Invited review

Humans on Earth: Global extents of anthropogenic land cover from remote sensing

CrossMark

Christopher Small*, Daniel Sousa

Lamont Doherty Earth Observatory, Columbia University, Palisades, NY 10964 USA

ARTICLE INFO

ABSTRACT

This review provides a perspective of the current state of the art in remote sensing of anthropogenic land cover and human-modified landscapes at global scales. The fact that humans have adapted to almost all of Earth's environments, yet remain strongly clustered within each of these environments influences both the nature of anthropogenic impact on Earth's landscapes and the challenges of mapping it. Remote sensing provides a consistent synoptic view of these environments by mapping the land cover associated with the anthropogenic land uses of settlement and food production, as well as their complement in forest cover. We give brief descriptions and illustrative comparisons of several current land cover products representing the global extents of settlements, agriculture and forests derived from remote sensing. To accommodate the challenges inherent to mapping any land cover at widely varying scales, we compare size distributions of spatially contiguous land cover (rather than total area) for several global land cover products. Despite the use of different sensors and different mapping criteria, there is remarkable consistency in the size distributions of these products – both within and across land cover class. Rank-size distributions of settlements, agricultural areas and forests are all well-described by power laws spanning more than four orders of magnitude in both area and number. This consistency in the form of the distributions suggests fundamental similarities among different types of land cover. The observed similarities can be explained by depicting land cover mosaics as co-evolving spatial networks sharing common processes of nucleation, growth and connection.

© 2016 Published by Elsevier Ltd.

Contents

---

1. Introduction

Over the past ∼60 years, humans' view of Earth has changed radically, from relatively narrow local perspectives based primarily on firsthand observation to broad global perspectives increasingly influenced by indirect sources of information. The parallel technological evolution of communication and imaging has contributed to this change. Improvements in imaging systems have generated an enormous volume of information not derived from direct experience. Advances in communications now allow us to disseminate, and consume, this information in volumes far exceeding those to which humans even one generation prior had been exposed. The perspectives of increasing numbers of people are influenced by remotely sensed images of Earth and the information derived from them. From meteorological satellites providing information about regional weather patterns to mobile sensors providing information about transportation network flows, remote sensing provides an increasing volume of indirect observation and perceived experience which is shaping the way humans view the world and their place in it. This review paper is an attempt to provide a perspective on the current state of the art in global land cover mapping that remote sensing provides on human-modified landscapes.

Human habitat on Earth is characterized by a duality in which humans have adapted to almost all of Earth's environments (Vitousek et al., 1997), yet remain strongly clustered within each of these environments. This duality influences both the nature of anthropogenic impact on Earth's landscapes (e.g. Ellis and Ramankutty, 2008) and the challenges of mapping it. Mapping the global distribution of human population within geophysical parameter spaces provides a simple way to quantify both aspects of this duality: the breadth of our adaptation across environments and the nature of the clustering (Cohen and Small, 1998; Small and Cohen, 2004). Fig. 1 illustrates the current distribution of human population and development relative to the climatic parameters of temperature and precipitation using a combination of population density from spatially gridded census enumerations (CIESIN, 2014) and economic development inferred from stable night lights imaged by satellites (Elvidge et al., 1999). Relative to mean annual temperature and

precipitation (New et al., 1999), it is apparent that both populated land area (as enumerated by census) and lighted human settlements span all major biomes on Earth (except ice sheets). However, the spatial Lorenz curve showing the cumulative distribution of 2010 population relative to cumulative enumerated land area (excludes Antarctica and parts of Greenland) shows that populations are strongly clustered in space at both national and global scales, with more than half of global population inhabiting less than 2% of enumerated land area. This also illustrates the importance of remote sensing for accurate quantification of human-modified landscapes. For reasons of logistics and cost, census enumerations and other survey-based metrics do not provide detailed depictions of dispersed rural populations, or even urban populations in many developing countries where accurate, detailed census are not conducted. In contrast, remote sensing provides uniform global coverage but introduces the challenge of interpretation.

Large area administrative units are used to represent dispersed populations in sparsely populated areas (most of the world) but inevitably lack spatial detail to explicitly show the distribution of the smaller settlements within the unit. However, remote sensing provides a wealth of information about these dispersed populations and their interactions with the landscape. The synoptic spatial coverage provided by satellite remote sensing is complementary to detailed *in situ* surveys that cannot provide complete spatial coverage at the scale of dispersed rural settlements.

By necessity, the scope of this review is limited to global products derived from satellite remote sensing. The vast majority of the remote sensing literature is devoted to development of algorithms and their implementation at local to regional scales. For reasons explained below, products and algorithms developed for specific local and regional applications do not generally scale to global extents. In the interest of representing the global extent of human modification of Earth's landscapes rather than the myriad facets of this modification, we focus on the relatively small number of global products that attempt to map different types of anthropogenic land cover consistently at global scales.

The structure of this paper is based on a sequence of background, description, synthesis and illustration. The
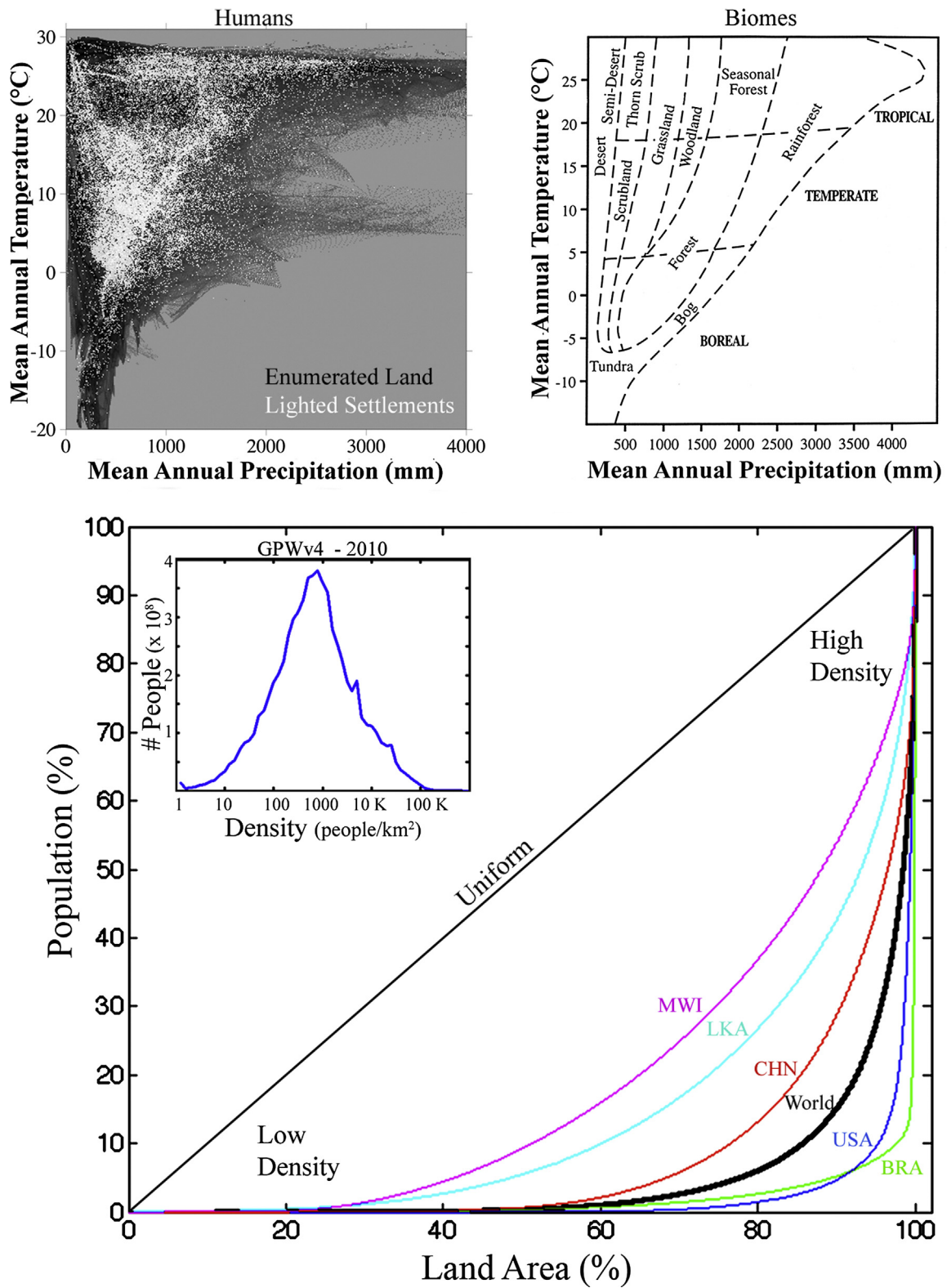
Fig. 1. Global dispersion and clustering of human population and settlements. Dispersion of population and lighted settlements over climatically defined biomes (top) reflects widespread human adaptation to terrestrial environments. Bivariate histogram (left) shows census-enumerated land area and lighted settlements in a climatic parameter space spanning major biomes (right - from Houghton et al., 1996). Clustering of population (bottom) reflects human tendency to congregate in settlements at global and national scales. Cumulative population on cumulative land area quantifies degree of spatial clustering in settlements. Malawi, Sri Lanka, China, USA and Brazil increase progressively in degree of relative clustering within their respective land areas. At census resolution more than half of the global population lives on less than 2% of the enumerated land area – but small settlements and dispersed populations are not resolved in large census units so this is a conservative estimate.

background section begins with a brief overview of the process of remote sensing and the data it provides. This overview is cast in the context of spatial, spectral and temporal resolution of measurements because these are the fundamental determinants of what can and cannot be quantified by remotely sensed observations. With the background of the observations established, we next introduce the two primary classes of products derived from remotely sensed data and the two classes of algorithms used to generate products from data. This is intended to provide a basis for understanding the strengths and limitations of the global products. The description section divides the currently available global products into the two most globally pervasive categories of anthropogenic land use: settlements and food production. Each product is described briefly using a common format summarizing the data, algorithms and primary characteristics of the product. In addition to the two primary categories of land use, we also include an overview of global forest cover products in the context of their use for monitoring deforestation. While the process of deforestation is often driven by the expansion of settlements and food production, other drivers also exist. We treat the mapping and monitoring of forests separately from the two primary categories of land use because the land use that follows deforestation cannot always be inferred unambiguously from the land cover that replaces the forest.

While the descriptors settlement and agriculture imply land use, most of these products infer land use from the land cover. Whereas forests represent a form of land cover with different uses (e.g. ecosystem services, silviculture, habitat). We make no attempt to impose rigid definitions on the products discussed. Rather we allow readers to infer land use or land cover specific to the product and its application. For each category, we show comparisons of products with comparable spatial resolutions as composite images. This is intended to illustrate both the similarities and differences among the products. For all three categories, we use the Ganges-Brahmaputra delta and its surrounding uplands and mountains for regional scale illustrations because it contains diverse examples of each land cover category, and because of our firsthand experience conducting fieldwork in all three categories throughout the delta over the past decade. Following the description of the global products, we present an illustrative synthesis of the global rank-size distributions of several of the products. The synthesis is not intended to provide a definitive accounting of the global extents but rather to illustrate some common characteristics of the distributions and scaling properties of each category of land cover. The scaling properties are relevant to the topic because they inform our understanding of the detection limits of the data and algorithms used to generate the products – and of the driving processes that are responsible for the current distribution of each land cover category. To provide a context for the scaling properties observed among the products, we introduce the concept of land cover as interspersed spatial networks. Treating land cover as spatial networks provides a basis for understanding changes in the distribution of land cover in the context of network growth or shrinkage. We follow the global comparisons with a discussion of the strengths and limitations of remote sensing and the principal challenges to mapping and monitoring anthropogenic land cover and its impact on the environment. Finally, to give readers a deeper understanding of the diversity and characteristics of the different types of land cover and land use around the world, we include a gallery of multi-scale satellite observations of a variety of human-modified landscapes in the form of an appendix (http://www.LDEO.columbia.edu/~small/LandCoverGallery). The gallery is central to the objectives of the paper because it provides the reader with numerous examples of the complexity of anthropogenic land cover and the challenges of mapping it consistently at global scales.

1.1. Background

1.1.1. Remote sensing, data and products

In light of our expectation that some readers may have little or no background or experience with remote sensing of land cover, we precede the review with a brief introduction to some key physical concepts of remote sensing. More in-depth discussions of these concepts and their importance for land cover mapping and monitoring are given by (Adams and Gillespie, 2006), (Jensen, 2007) and (Small, 2009) and the references provided therein.

For the purposes of this discussion, we distinguish between the data derived from remotely sensed measurements and the products derived from the data. There is a continuum of processing stages between the initial measurements collected by a sensor and the final products interpreted by users. Each stage of processing employs algorithms that combine observations with assumptions to produce its respective product. Here we consider *data* to be measurements processed with standardized algorithms intended for generalized application in which the assumptions are not specific to the applications. This includes processes like geometric rectification and radiometric calibration. We consider *products* to be outputs derived from data for which the assumptions built into the processing algorithm are not standardized and may vary considerably from one product to another. As a result, different products designed to map the same nominal land cover class may use the same input data but produce very different outputs.

1.2. Resolution; spatial, spectral, temporal

The limits of what an imaging sensor can measure are determined by the combination of its spatial, spectral and temporal resolution. The resolving power of an instrument refers specifically to the smallest interval quantity that it can measure. In the case of spectral resolution the interval is a range of wavelengths of radiation, whereas in the case of spatial resolution it is a distance. Temporal resolution is determined by the revisit time between image acquisitions. Differences in spatial and spectral resolution are illustrated in Figs. 2 and 3 (explained in detail below). Fig. 4 uses a time-space cube of multitemporal imagery to illustrate the temporal variability of land cover and the importance of temporal sampling to supplement the interpretation of individual images.

1.2.1. Spatial resolution

The spatial resolution of an imaging sensor determines the size of the smallest object it can image coherently. The spatial resolution is dictated by the angular Instantaneous Field Of View (IFOV) and its corresponding ground sample distance (GSD) that determines the size of a single picture element (pixel). A sensor cannot produce a coherent image of an object smaller than a single pixel – although sensors can often detect the presence of objects smaller than a single pixel as they contribute to the aggregate measurement of the pixel. Because the sensor measures a single spatial average of the radiance reflected from all illuminated objects within its IFOV, the reflectance of each of these objects contributes to the aggregate radiance measured by the sensor. This process is referred to as spectral mixing. If the different reflectances that contribute to this aggregate radiance are known, or can be inferred, it is sometimes possible to estimate the relative abundance of the different objects within the IFOV. This is generally referred to as spectral unmixing.

To avoid confusion resulting from subjective terms like high, medium and low, we specify spatial resolution scales explicitly using the metric prefixes sub-meter ($\lambda < 1.0$ m), meter ($1.0 \leq \lambda < 10$ m), decameter ($10 \leq \lambda < 100$ m), hectometer ($100 \leq \lambda < 1000$ m) and kilometer ($1.0 \leq \lambda < 10$ km).

1.2.2. Spectral resolution

Optical sensors on satellites are conceptually similar to sensors used in digital cameras. Whereas camera sensors image brightness of visible light in red, green and blue wavelength bands, sensors used on satellites also image reflected sunlight in infrared spectral bands. Extending the concept of color from visible into infrared wavelengths provides a new view of the world because many materials that are indistinguishable in the visible have distinct absorptions in the infrared. The spectral resolution and dynamic range of a sensor determine how many and which colors it can distinguish. This is important because land cover can sometimes be discriminated on the basis of color. The analogous concept for active microwave sensors is more complex but related to the relationship between the wavelength of radar used and the presence of physical structure (e.g. roughness) at the scale of that wavelength. This is why the radar backscatter response of a target (e.g. forest) depends strongly on the wavelength of radar used to image it.

1.2.3. Temporal resolution

Temporal resolution refers to the temporal frequency of sampling by repeat imaging. Temporal resolution is of monumental importance for both mapping and monitoring land cover because of the variety of processes that cause landscapes to change on a range of different time scales. In order to distinguish actual land cover change from apparent changes related to atmospheric conditions, illumination, soil moisture and vegetation phenology, it is critical to understand how these sources of apparent change are manifest, as well as the actual changes being monitored. In the past, temporal resolution was determined by the revisit frequency of a specific satellite orbit and the extent of adjacent swath overlap. Increasingly, temporal resolution also depends on the number of satellites with comparable sensors operating simultaneously in the same constellation.

1.3. Sensors and satellite missions

Imagery collected by satellites is commonly referred to either by the name of the sensor that makes the measurements or by the name of the satellite that carries the sensor. Strictly speaking, referring to the sensor is more precise since the same satellite may carry multiple sensors measuring different quantities. For consistency, we use the name most commonly used in the literature, which is generally the sensor on multi-instrument platforms (e.g. MODIS) and the satellite when it carries only one sensor or when one of multiple sensors accounts for the vast majority of usage (e.g. Landsat).

MODIS—The Moderate Resolution Imaging Spectroradiometer is a NASA sensor designed to image the earth in 36 spectral bands at optical and thermal IR wavelengths (Barnes et al., 1998). The MODIS sensor is operational aboard two NASA satellites, Terra and Aqua, in sun-synchronous near-polar orbits. The seven spectral bands used for the land cover products in this study are two 250 m resolution bands imaging in the red and near infrared and five 500 m resolution bands imaging in the blue, green, and three bands of the shortwave infrared. Six of these seven bands are similar to optical bands imaged by the TM/ETM+/OLI sensors on the Landsat satellites. The combination of the 2330 km swath and placement on two satellites provides for complete Earth coverage every one to two days. The Terra satellite was launched December 18, 1999 and Aqua was launched May 4, 2002. Despite the MODIS sensor design life of six years, both MODIS instruments are still collecting usable land surface data. All MODIS data are publicly available. Additional information about MODIS is available at: http://modis.gsfc.nasa.gov/about/

MERIS—The Medium Resolution Imaging Spectrometer (MERIS) was an instrument aboard the Envisat satellite in a sun-synchronous polar orbit (Rast et al., 1999). Envisat was built and operated by the European Space Agency (ESA). MERIS collected data in 15 spectral bands in the visible and near infrared spectrum. MERIS imaged a 1150 km swath, resulting in complete global coverage every 3 days. Envisat was launched in March 2002. Communication was lost with Envisat in April 2012. The designed life of Envisat was 5 years. Additional information about MERIS is available at: https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/envisat/instruments/meris.

Landsat—The Landsat program provides the longest public continuous record of Earth imaging from space, operating from the launch of Landsat 1 on July 23, 1972 to the present. The products described here rely on the most recent sensors used in the Landsat program: the Thematic Mapper (TM, Landsat 4 and 5), Enhanced Thematic Mapper Plus (ETM+, Landsat 7), and Operational Land Imager (OLI, Landsat 8). The primary land surface data collected by all of these sensors are 6-band, 30 m resolution imagery collected in the blue, green, red, near infrared, and two shortwave infrared bands in a 185 km swath. Landsat revisit time is 16 days. Landsat 5 was launched in March 1984 and collected usable imagery until being decommissioned in 2013. Landsat 7 was launched in April 1999 and continues to operate. On May 31, 2003 the Scan Line Corrector on Landsat 7 failed, resulting in scan line gaps in all ETM+ imagery collected since that date. Landsat 8 was launched in February 2013 and is currently fully operational. For a history and prospectus of the Landsat program see (Loveland and Dwyer, 2012). Following the opening of the Landsat archive in 2008, all Landsat data are available publicly online from the USGS. Additional information about the Landsat program is available at: http://landsat.usgs.gov/about_project_descriptions.php

Sentinel—The Sentinel series of satellites is planned to provide a 6-component suite of Earth observations. As of the time of writing, the first two of these components, Sentinel 1-A and 2-A are currently operational. Sentinel 1-A was launched in April of 2014 and Sentinel 2-A was launched in June of 2015. Sentinel 1 and Sentinel 2 are each based on pairs on twin satellites (A and B) orbiting the earth in parallel. Sentinel 1 features a Synthetic Aperture Radar (SAR) sensor operating in the C-band. Sentinel 1 is designed to collect data in 4 modes, with a range of polarizations (quad or dual), spatial resolutions (from 50 m to 9 m), and swath widths (400 km to 20 km). The Sentinel 1 constellation is designed to have a 6 day revisit time when both satellites are operational. Sentinel 2 features the MultiSpectral Instrument (MSI) sensor, collecting data in 13 spectral bands in the visible and infrared portions of the electromagnetic spectrum. Each band is collected at 10 m, 20 m, or 60 m resolution, depending on wavelength. Sentinel 2 images in a 290 km swath and the constellation is designed to have a 5-day revisit time. All Sentinel data are planned to be available publicly online from the ESA. Additional information about the Sentinel program is available at: http://sentinels.copernicus.eu

Commercial Sensors—IKONOS, the first meter-resolution commercial earth imaging satellite, was launched on September 24, 1999 by the Lockheed Martin Corporation and subsequently operated by the Space Imaging Corporation. The IKONOS sensor collected an 11 km wide swath of multispectral VNIR (blue, green, red, NIR) imagery at 4 m and simultaneous panchromatic imagery at 1 m spatial resolution. IKONOS was decommissioned in April 2015. Since the launch of IKONOS, the Space Imaging, Digital Globe, GeoEye, and Astrium corporations have launched and operated similar satellite systems with meter resolution sensors imaging at

VNIR wavelengths. Digital Globe operates the Quickbird (now decommissioned), GeoEye and WorldView-2 and 3 satellites and Astrium operates the Pléiades constellation – all carrying VNIR imaging sensors with 2-3 m multispectral and 0.5 m panchromatic bands.

TanDEM-X—TerraSAR-X add-on for Digital Elevation Measurements TanDEM-X is a satellite radar interferometer produced by two TerraSAR-X satellites orbiting close together to form a larger aperture than would be feasible from a single spacecraft. The TerraSAR-X satellites, built and operated by the Deutsche Zentrum für Luft- und Raumfahrt (DLR), use X band microwave radar to form a synthetic aperture capable of imaging surface elevations at 12 m resolution (Krieger et al., 2007). Compared to optical systems, the primary benefit of radar is the ability to image at night and through clouds. An additional benefit for mapping built environments is related to the tendency of vertical walls and the adjacent ground surface to form corner reflectors which backscatter radar waves strongly in the direction from which they are received – thereby causing areas with a high density of vertical walls to appear much brighter than other areas regardless of imaging geometry. While the primary objective of the TanDEM-X mission was to map global land surface elevations at higher spatial resolutions than previously possible, the resulting measurements are also able to resolve spatial textures associated with buildings and other vertical structures. This, combined with the enhanced backscatter from vertical structures and land cover-specific speckle characteristics, allows TanDEM-X data to be used to discriminate features associated with built environments (Esch et al., 2011). Additional information about TanDEM-X is available at: http://www.dlr.de/hr/desktopdefault.aspx/tabid-2317/3669_read-5488/.

DMSP-OLS—Defense Meteorological Satellite Program – Operational Line Scanner The DMSP series of satellites have been collecting night time imagery of the Earth surface since the 1970's. Although originally collected on film, the OLS sensors have recorded digital measurements of night light since 1992. The Earth Observation Group at the National Geophysical Data Center has produced and disseminated annual night light composites since 1994. The OLS sensor images emitted visible and thermal infrared radiance in a 3000 km swath twice per night. Spatial resolution varies from nadir (directly beneath the sensor) to swath edges so multiple images are composited at a grid resolution of 30 arc seconds (~1 km at Equator). Although the sensor was designed to image moonlit clouds, for most research applications images are used from early evening (7–10 pm) acquisitions on non-moonlit nights. More detailed descriptions of the data, products and applications are given by (Elvidge et al., 1997, 2001). Since 1992, a series of five DMSP missions have provided continuous imaging of night lights. Intercalibration across OLS sensors on different satellites provides a 22 year time series of annual global composites of stable night light. Annual composites and individual acquisitions of night lights have been used for a wide range of monitoring applications ranging from gas flaring (Elvidge et al., 2009) to refugee camps (Li and Li 2014).

VIIRS—Visible Infrared Imaging Radiometer Suite The VIIRS sensor was launched on board the NASA-NOAA Suomi satellite in 2011. The day/night band (DNB) collects low light imagery in a 3000 km swath at a resolution of ~500 m at nadir with an equator overpass time of ~1 AM. Like OLS night light measurements, individual VIIRS acquisitions are typically composited to reduce cloud contamination and temporally intermittent light sources like fires. Unlike OLS, VIIRS provides higher dynamic range, on board calibration and multiple optical bands useful for discrimination of different light sources. More detailed descriptions of the data, products and applications of VIIRS imagery are given by (Elvidge et al., 2013; Miller et al., 2013). In addition to resolving smaller lighted settlements than OLS, VIIRS has sufficient spatial resolution and low light imaging ability to individual light sources like fishing boats (Elvidge et al., 2015).

1.4. Satellite image interpretation

1.4.1. Non-uniqueness of color

The principal advantage of multispectral sensors is their sensitivity to the variations in surface reflectance that give rise to the phenomenon that we perceive as color – for both visible and infrared wavelengths of light. The principal limitation of multispectral sensors is related to the non-uniqueness of color. While color may be a valuable discriminant in many situations, it is often not diagnostic by itself. The non-uniqueness of color presents a challenge because different classes of land cover can have indistinguishable colors while the same class of land cover can have different colors. This inconvenient truth can have grim implications if assumptions made about land cover reflectance do not hold true in different environments. This is one of the primary challenges of mapping land cover consistently at global scales. Land cover spectral characteristics that may be robust in some environments are often not valid in other environments.

1.4.2. Importance of scale

The concept of spatial scale is fundamental to remote sensing of any type of land cover (Woodcock and Strahler, 1986). The information provided by remote sensing depends critically on two related spatial scales. The spatial scale of the sensor IFOV determines what contributes to the radiance measurement associated with each pixel spectrum. The spatial scales of the individual objects (e.g. roofs, streets, trees, agricultural fields, etc) in the land cover mosaic determine the extent of spectral mixing that occurs within the sensor IFOV that images them. Objects that are much larger than the pixel are spatially oversampled and represented as coherent features in the image. Objects that are comparable to, or smaller than, the size of the pixel are not spatially resolved but do contribute reflected radiance to the aggregate pixel measurement in proportion to their areal abundance. This allows spectral unmixing to be used to estimate relative proportions of land cover at sub-pixel scales (Adams and Gillespie, 2006; Adams et al., 1995, 1986; Smith et al., 1990). This relationship between resolution and scale is particularly important in remote sensing of any environment where detection of individual objects is important. For example, Welch (1980) recognized systematic variations in the size of buildings and streets in many urban environments and proposed that a minimum spatial resolution of 5 m is generally required to discriminate individual components of the urban mosaic.

With these caveats in mind, we illustrate the combined effects of scale and reflectance on imaging of different landscapes with a set of multi-scale illustrations in Fig. 2 and in the Gallery at the end of the paper. Each figure shows each image at full sensor resolution, in which each pixel in the digital version of the figure corresponds to a sensor pixel. Enlarging any panel in Fig. 2 to full resolution will show the resolution limit of the sensor image used. The meter scale examples were collected by the commercial sensors Ikonos (4 m), Quickbird (2.8 m), GeoEye (2.0 m) and WorldView-2 (2.0 m). All show visible red and blue spectral bands in the corresponding channel of the image, but show the near infrared (NIR) band in the green channel to distinguish NIR-bright vegetation from shadows, water and other dark surfaces. The center panel in each example illustrates the 30 m resolution and swath width of the sensors carried on the Landsat 4, 5, 7 and 8 satellites. Each Landsat example uses the same blue, NIR and red bands as the corresponding higher resolution example and shows exactly the same location acquired as close as possible in time

Fig. 2. (a) Multiscale comparison of urban land cover at 4.0 m (top), 30 m (middle) and 250 m (bottom) pixel resolution. Enlarge to see full resolution limit. Ikonos © 2016 Digital Globe Inc. (b) Multiscale comparison of rural land cover at 2.8 m (top), 30 m (middle) and 250 m (bottom) pixel resolution. En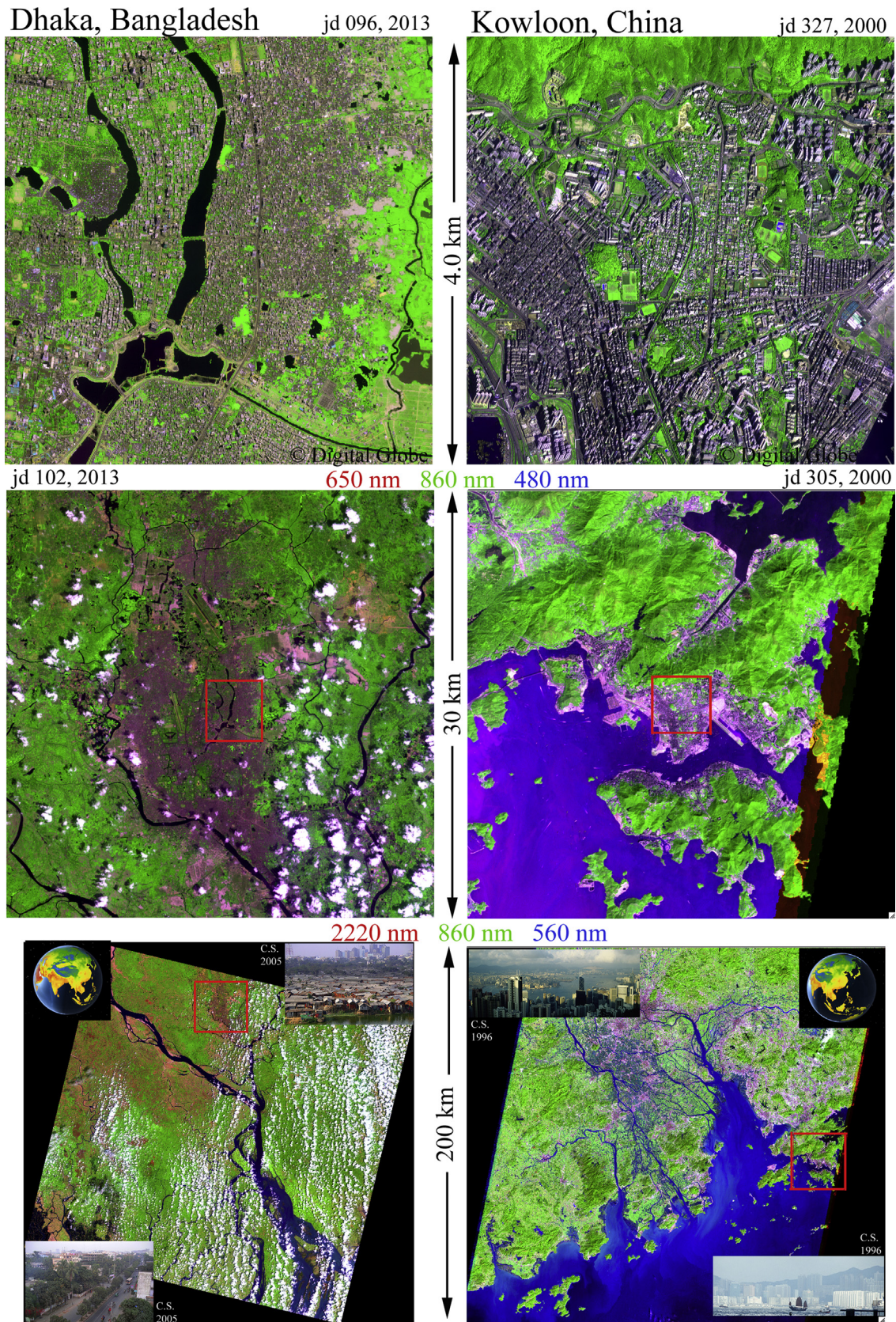large to see full resolution limit. Quickbird © 2016 Digital Globe Inc. (c) Multiscale comparison of forest cover at 2.0 m (top), 30 m (middle) and 250 m (bottom) pixel resolution. Enlarge to see full resolution limit. WorldView-2 © 2016 Digital Globe Inc.
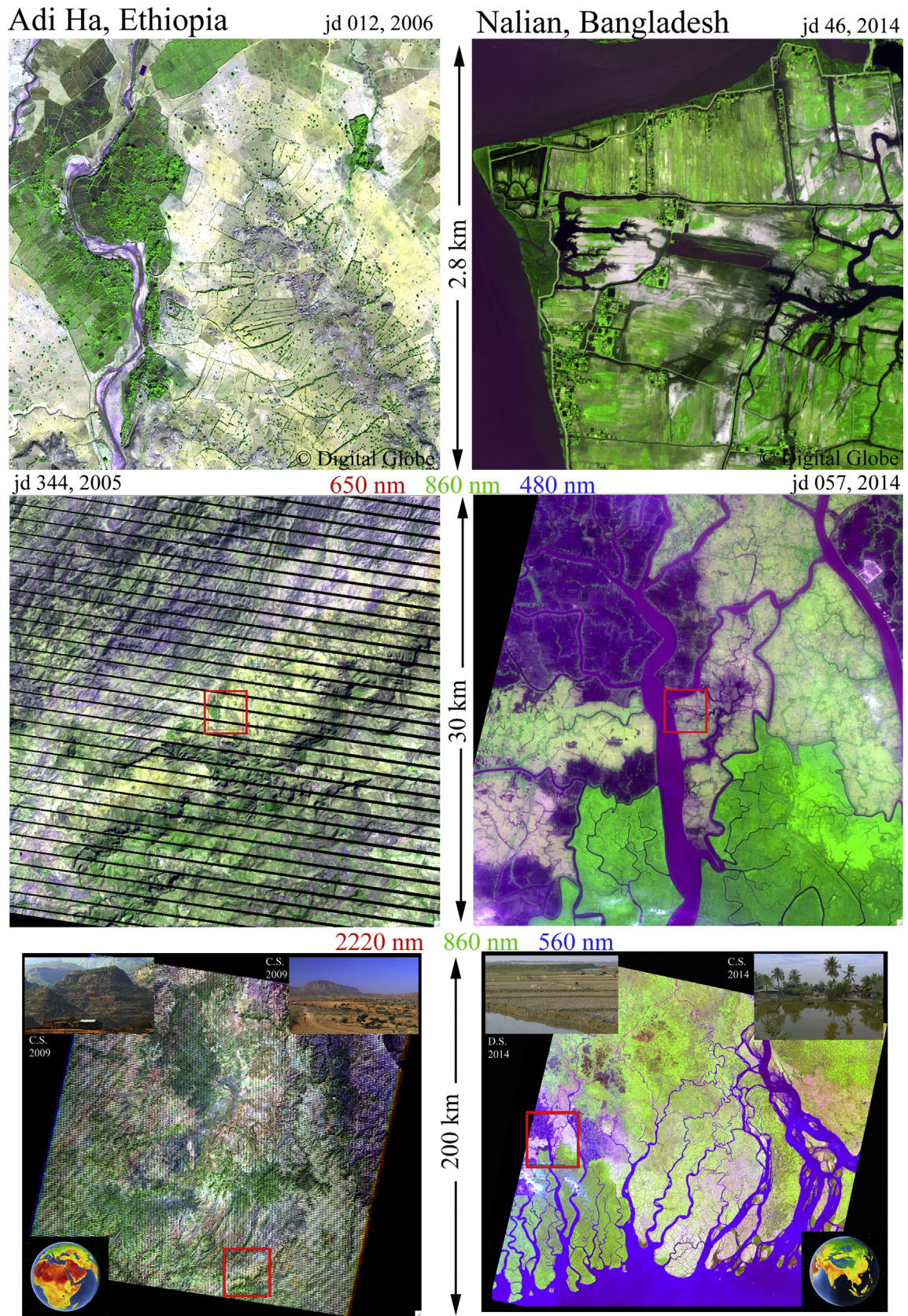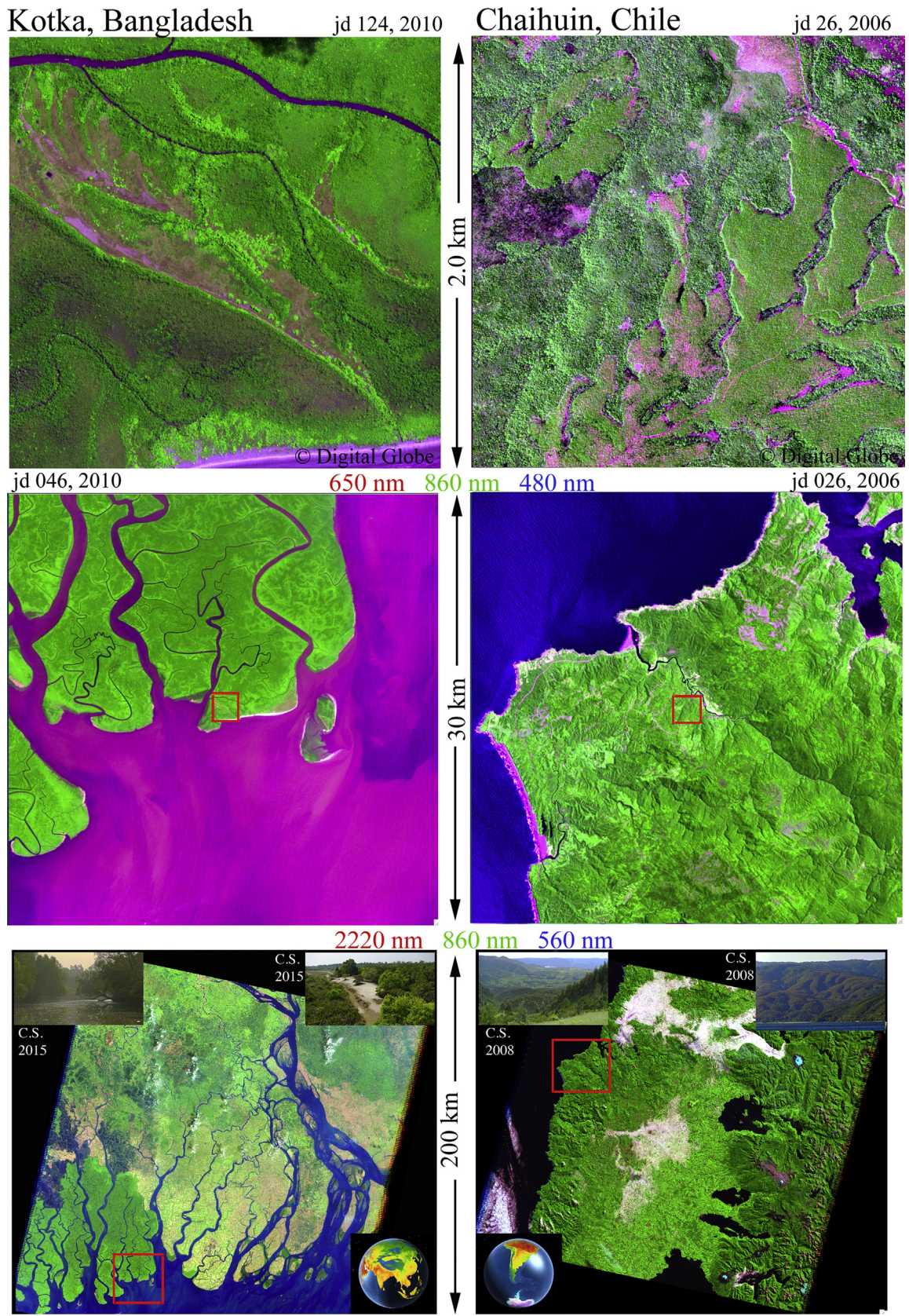
Fig. 2. (Continued)

Fig. 2. (Continued)

(Julian day). The bottom panel in each example shows the entire Landsat scene, but as a composite of visible, NIR and SWIR bands resampled to match the 250 m GSD of MODIS. Showing the same landscape at three different spatial scales and resolutions, in different parts of the optical spectrum, is intended to convey some of the spatial and spectral complexity of different landscapes at different scales.

To illustrate the diversity of land cover in different environments Fig. 2 shows three pairs of multi-scale examples: densely populated urban environments, sparsely populated rural environments and effectively unpopulated forest environments. In the urban example (Fig. 2a), both cities contain a variety of building types and densities, and varying amounts of vegetation and open space. The contrast between the large residential skyscrapers throughout Kowloon and the much smaller buildings in Dhaka and central Kowloon gives some indication of the scale range of individual buildings. In the neighborhood of Korail, the large dark gray area north and west of the center of the Dhaka image, the size of individual buildings, and the separation between buildings, is generally below the 4 m resolution of the image. A more detailed view of Korail, and the challenges of mapping informal settlements, are given by Stevenson et al. (2013) at: http://blog.brac.net/2013/09/frugal-map-making-experiences-from-korail/. In both cities, the Landsat image reveals the complexity of land cover both within and at the periphery of the densely built up urban core – despite the high contrast between the built surfaces (buildings and roads) with the surrounding vegetation (agricultural and native) and water. The full Landsat scenes show the corresponding complexity of the regional land cover mosaic that each city evolves within, and how it would be seen by wide swath sensors like MODIS and MERIS.

In contrast, the landscapes shown in Fig. 2b illustrate dispersed rural populations and their landscape modifications, primarily in the form of agriculture. Even in cases where the agricultural fields are fallow, or still at early stages of cropping, sharp contrasts in reflectance are apparent at edges of plots. These variations in color may be due to differences in either soil moisture content or tillage (i.e. roughness), or both. In both cases, the scale of individual dwellings approaches the resolution of the sensor and are apparent only when the image is enlarged. Identification of dwellings is further complicated in the Bangladesh example because almost all are built within clusters of trees which can obscure the buildings beneath. In both examples, dwellings are often roofed with non-photosynthetic vegetation (e.g. thatch) and natural materials that may be compositionally indistinguishable from other natural materials found throughout the landscape. At the decameter scales of the Landsat images, the dispersed settlements are indistinguishable – although abrupt transitions in land cover are still apparent as sharp boundaries in reflectance. In the arid environment of northern Ethiopia, sharp boundaries in reflectance can also correspond to transitions in substrate reflectance related to underlying geologic structure. In the deltaic environment of southern Bangladesh the sharp boundaries between bright and dark land cover generally correspond to transitions between fallow agricultural fields and large shallow ponds used for aquaculture.

Different types of forest cover are illustrated at multiple scales in Fig. 2c. The example from Bangladesh shows a transitional area in the Sundarban mangrove forest on the southern coast of the delta. In addition to the mud-dominated intertidal environment of the mangrove forests in the northern part of the meter resolution image, we show transitions to more elevated sand-dominated storm deposits and the dry forests that occur on these soils. The example from southern Chile shows diverse native forest interspersed with managed pine and eucalyptus forests at varying stages of growth. Whereas the spatial variability in canopy closure (seen as variations in brightness of green) in the Bangladeshi

mangrove example is entirely natural and related to age and species composition, variation within the Chilean montane forest is due to both natural species variation and varying stages of silviculture conducted throughout the region.

Limitations imposed by spatial resolution and spectral non-uniqueness have important implications for our ability to identify and map land cover accurately in different settings. Different land cover types are interspersed at different spatial scales in different landscapes. Our ability to distinguish them depends on the spatial scale at which they are interspersed relative to the resolution of the sensor. The finer the scale at which the land cover types are interspersed, the higher the spatial resolution required to distinguish them. This is illustrated with a mixture scale tetrahedron in Fig. 3.

The base of the tetrahedron is a simple ternary diagram showing the relative areal abundance of three broad categories of land cover associated with human settlements (e.g. buildings and roads), food production (e.g. agricultural fields, pasture) and the background of "natural" land cover on which they are superimposed. Lacking a more general concise term, we use natural as a variable to describe whatever land cover type complements the land cover associated with settlements and food production. Different landscapes occupy different parts of the ternary space, depending on the relative abundance of each land cover type at the scale of analysis. This ternary space is completed with an additional dimension representing the spatial scale at which the different land cover types are interspersed. This scale is a necessary complement to the relative areal abundance because it determines both the nature of the interface between the land cover types and our ability to distinguish them with a sensor of a given spatial resolution. The nature of the interface is a fundamental determinant of the nature of interactions between the land covers and their associated land uses. For example, flows of material and energy between settlements, food production areas and natural areas can be strongly influenced by the scale and fractal dimension of the boundary separating them. Fig. 3 illustrates this concept using nine examples of different landscapes containing different areal abundances of each land cover type and different scales of interface separating them.

1.4.3. Importance of temporal resolution

Although often not considered in studies using single date image acquisitions, temporal resolution is of monumental importance for characterization of most landscapes. Daily changes in atmospheric opacity, weekly changes in soil moisture, monthly changes in illumination and seasonal changes in vegetation phenology all result in apparent and actual changes in the way sensors see landscapes and the way we interpret the sensors' measurements. Temporal resolution is particularly important for mapping land cover that can be distinguished by seasonal variations in vegetation phenology and illumination. Fig. 4 attempts to demonstrate the importance of temporal resolution by showing a time-space cube of vegetation index maps derived from 15 years of biweekly MODIS composites of the Ganges-Brahmaputra delta. The sides of the cube show seasonal variations in vegetation abundance of the pixels on the corresponding edges of the cube. Temporal patterns of single and multiple cycles per year reveal a diversity of agricultural patterns related to different stages of multiple cropping. The annual monsoon cycle is also prominent in both natural and cultivated phenologies. The front of the cube shows a phenology map derived from the spatiotemporal patterns of all the pixel time series in the cube using a temporal mixture model. The colors on the front of the cube correspond to the different phenology patterns of the temporal endmembers shown below the cube. Spatio-temporal characterization of the cube identifies these three endmember time series as the most
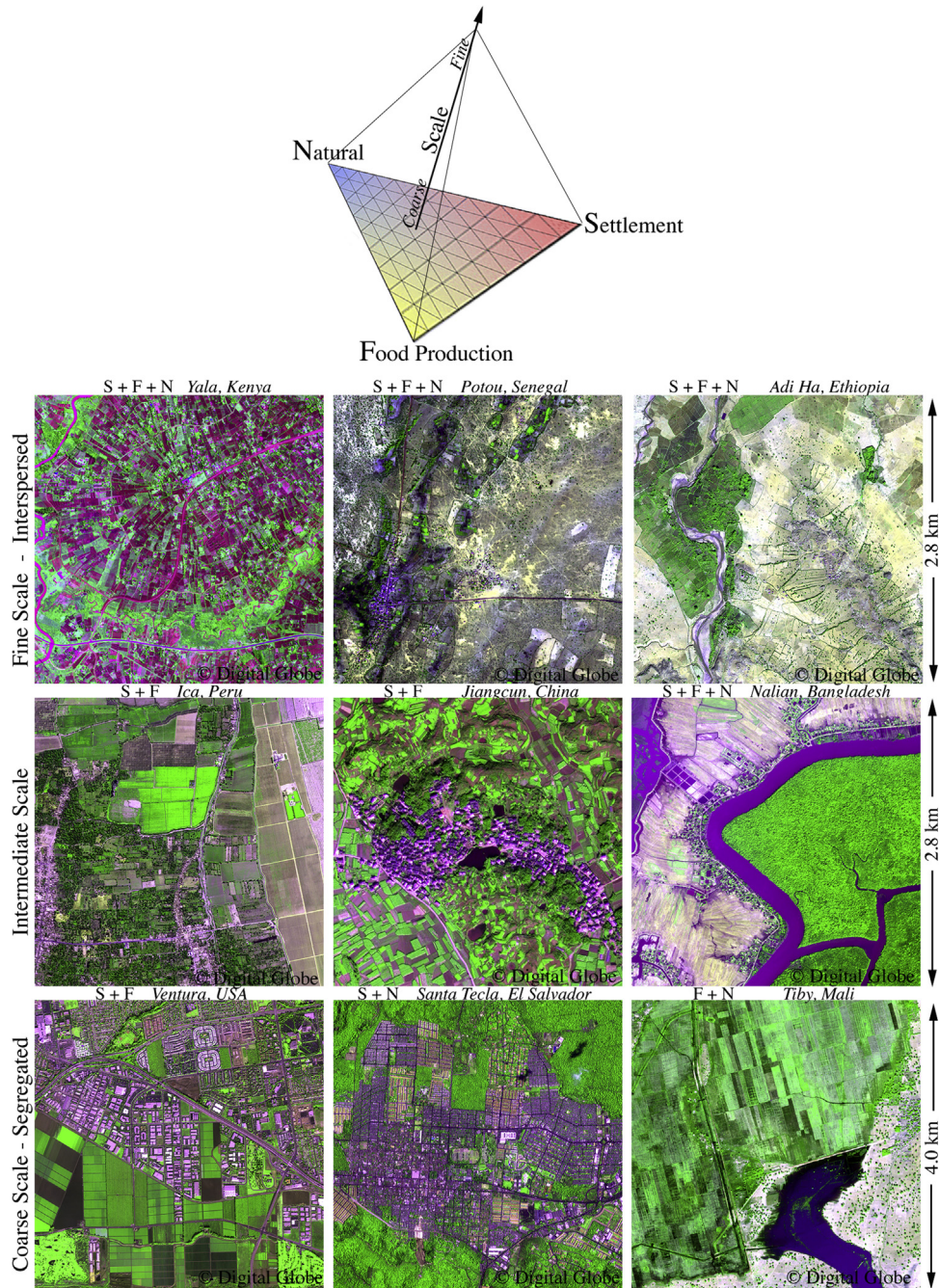
Fig. 3. Spatial scale and land cover heterogeneity. Ternary distributions of land cover can represent a continuum of mixtures at multiple scales. As land cover becomes increasingly interspersed and the characteristic scale of individual units becomes finer, the interface between units becomes increasingly fractal and discrimination becomes more difficult. Ikonos © 2016 Digital Globe Inc.

distinct of all in the cube (Small, 2012). Temporal mixture modeling provides continuous estimates of the degree of similarity of each pixel time series in the cube to each of the three phenological endmember time series. Despite the continuous model used to map the phenologies, sharp transitions are observed throughout the image. Different multi-cycle phenologies correspond to different planting cycles, often related to timing of annual monsoon flooding.

Annual cycles in upland forests result from both deciduous and evergreen vegetation under varying seasonal illumination and monsoon rainfall. Even in the Sundarban mangrove forest, evergreen vegetation shows some seasonal variation – although

the lower degree of color saturation indicates that none of the periodic phenologies is as predominant as in other parts of the map. Knowledge of these seasonal cycles, and interannual changes, is essential for characterization of vegetated land cover because different types of vegetation can be indistinguishable at certain times of the year but distinguished by differences in temporal cyclicity if at least three images per year are used in the analysis. Even in built environments with little vegetation, seasonal variations in illumination and dust cover cause changes in aggregate reflectance at decameter scales. Fig. 5 shows the contrast in New Delhi and surroundings between wet and dry season of the same year. Notice that during the dry season the reflectance of the densely built area is very similar
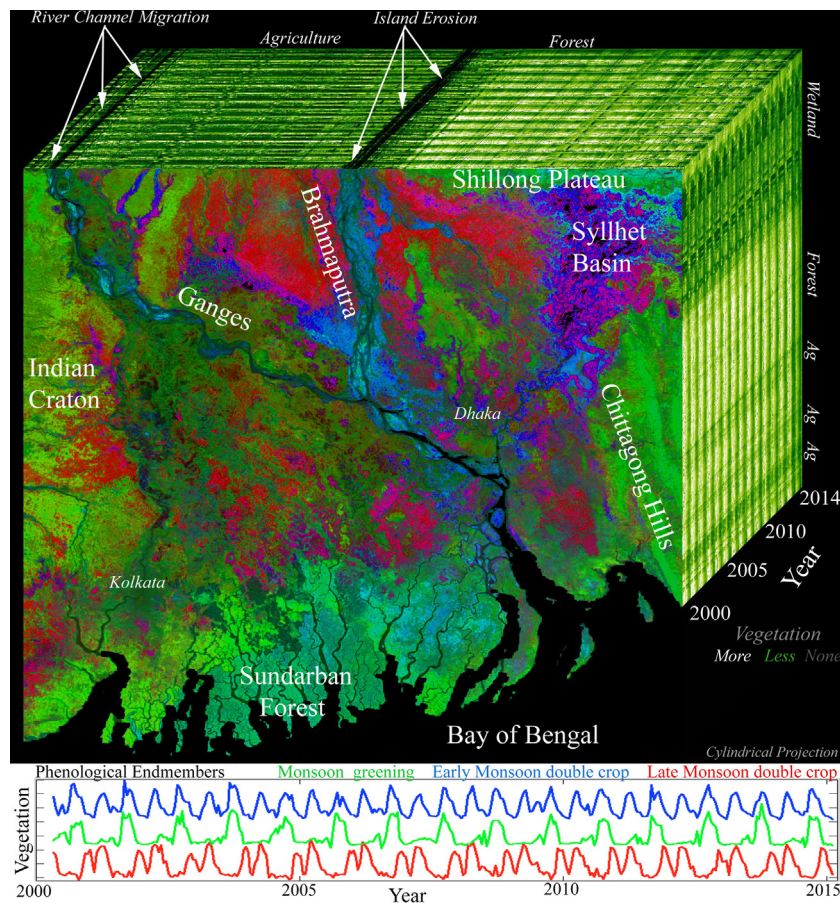
Fig. 4. Time-space cube showing MODIS vegetation index time series of the Ganges-Brahmaputra delta. Front face is a continuous field vegetation phenology map showing spatial gradations among three endmember phenologies plotted below. Sides of cube show 15 years of biweekly vegetation index time series for pixels at edges of cube. Spatial variations of vegetation phenology reflect different multi-crop agricultural cycles and annual monsoon cycle.

to that of bare soil in the surrounding agricultural areas. These illustrate the kind of changes that can confound classification of different land cover types, especially when based on single-date imagery.

1.5. Models and products; continuous and discrete

In the context of this review, we consider two types of products: continuous physical fields and discrete thematic classifications. These differ primarily in their depiction of the landscape. Discrete thematic classifications assume that the reflectance of each pixel uniquely identifies that pixel as a member of one, and only one, thematic class. Thematic class definitions are often based on inferred land use as well as inferred land cover. Thematic classification implicitly assumes that the measured spectral properties of each land cover type uniquely distinguish it from other types. In contrast, continuous fields require no assumptions about uniqueness of physical properties or homogeneity of land cover within a pixel. Because continuous fields can represent subpixel variations in relative abundance of different land cover types, they can accommodate the process of spectral mixing and the existence of mixed pixels – whereas mixed pixels violate the principal assumption of homogeneity upon which discrete classification depends. As a result, discrete classifications can represent land cover accurately when the scale of the pixel is much finer than the scale of the land cover units, but tend to be error-prone when the scale of the pixel approaches the scale of the land

cover unit. While any continuous field can represent discrete (abrupt) transitions in land cover as steep gradients in fractional abundance, no discrete classification can represent gradational transitions in land cover accurately. The principal advantage of discrete classifications is related to their conceptual similarity to traditional maps with distinct boundaries between land cover units or land use types. The principal advantage of continuous fields is their ability to represent spatial gradients in land cover and subpixel heterogeneity across a wide range of scales.

1.5.1. Continuous fields and mixture models

Land cover can be represented as continuous fields of its constituent components (e.g. vegetation, soil, rock, water) at a range of spatial scales. The relative amount of each land cover component within a pixel determines the aggregate response measured by the sensor. In this sense, a landscape can be represented as a combination of continuous fields of relative abundance of each land cover component. Continuous fields of land cover components are especially convenient because they accommodate the fixed spatial scale imposed by the IFOV of sensors. Areal abundance of land cover components (e.g. vegetation, water, soil) at the scale of individual pixels provides a convenient and easily measurable basis for representing a variety of land cover types (e.g. forest, agriculture, wetland) from which land use may be inferred. At the scale of the pixel, estimates of land cover abundance are generally derived using a linear spectral mixture model in which the pixel is represented as a spatial mixture of spectral endmember abundances. The principal
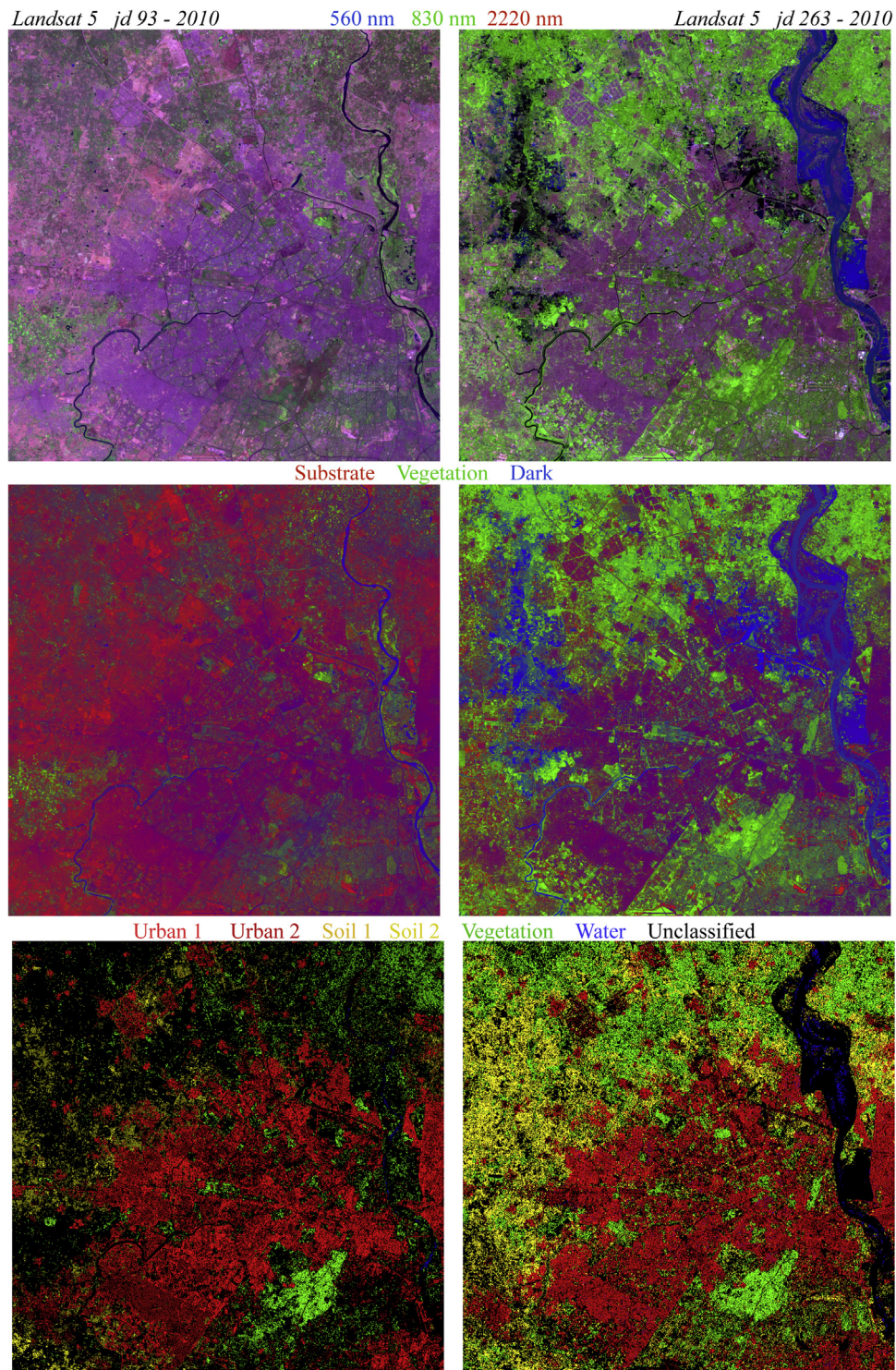
Landsat 5   jd 93 - 2010   560 nm 830 nm 2220 nm   *Landsat 5   jd 263 - 2010*

Substrate   Vegetation   Dark

Urban 1   Urban 2   Soil 1   Soil 2   Vegetation   Water   Unclassified

Fig. 5. New Delhi Landsat classification. False color composites (top) using identical enhancements of calibrated imagery show seasonal variability of land cover reflectance due to changes in vegetation, soil moisture and illumination. Continuous land cover fractions (middle) derived from reflectance accommodate subpixel mixing of land cover. Discrete thematic classifications (bottom) of each image using the same training sites differ considerably because the class definitions cannot accommodate seasonal changes in land surface properties.

challenge of unmixing mixed pixels is related to the choice and number of endmembers used in the model. An excellent treatment of the use of continuous fields and spectral mixture models to represent landscapes is given by (Adams and Gillespie 2006). Fig. 5 (center) shows the seasonal variations in land cover in and around New Delhi represented as subpixel fractions of rock and soil substrate, vegetation and dark surfaces like water and shadow. The fractions were derived from an inversion of a spectral mixture model of Landsat spectra that is conceptually analogous to the temporal mixture model used to produce the phenology map in Fig. 4.

1.5.2. Discrete thematic classification; supervised and unsupervised

Discrete thematic classification is the most commonly used approach to derive maps from multispectral imagery. Thematic classifications use the distribution of pixels in multidimensional feature spaces as the basis for decision rules to assign every pixel in an image membership in one, and only one, thematic class (urban, forest, water, agriculture, etc). The implicit assumption in discrete thematic classification is that the pixels cluster within the feature space and that decision rules can be found to separate pixels of different clusters (classes) accurately. Thematic classifications use classes corresponding to thematic characterization of land use – as inferred from land cover. The assumptions on which the inference is based influence the accuracy of the resulting classification.

Discrete classifications can be categorized as supervised or unsupervised. In both cases, decision boundaries segregate the feature space to define discrete classes. In the case of supervised classification, the decision boundaries are derived from statistical properties of training samples of pixels selected for each class. The assumption is that the training samples are sufficiently representative of each class to define decision boundaries that consistently separate pixels into distinct classes. In the case of unsupervised classification, no training samples are provided and the feature space is segregated by decision boundaries on the basis of the distribution of all pixels within the feature space. The assumption is that the distribution of pixels is sufficiently clustered within the feature space to allow an algorithm to correctly identify decision boundaries that naturally segregate the feature space into physically meaningful classes without prior identification of the classes. While no training samples are provided in unsupervised classification, several parameters (e.g. number of classes, separability, homogeneity) must be specified a priori. The choice of parameters used in the classification can have a profound influence on the resulting map. Uniqueness and sensitivity of results are important aspects of classifications – particularly when performed by algorithms with multiple degrees of freedom.

Ultimately, the accuracy of thematic classification of objects coarser than the scale of the pixel depends on the criteria chosen to define the classes. If the criteria are sufficient to distinguish between the specified classes in all instances (spatial and temporal), then the classification should produce accurate results. However, insufficient criteria in a single class can corrupt other classes and criteria that produce accurate results in some instances may not do so in other instances. Selection of sufficiently general criteria is one of the principal challenges of developing robust classifications at global scales.

An important benefit of thematic classification is its consistency with the simplistic depictions of land use that maps have used throughout history. An important limitation of thematic classification is the inaccuracy introduced by oversimplification of the continuous, multiscale nature of many landscapes. In comparison, an important limitation of continuous field depictions of landscapes is the requirement that the user expand their conception of landscapes as interspersed, multiscale mosaics of land cover with potentially complex relationships to land use. However, an important benefit of continuous field depictions of landscapes is the ability to represent their true complexity in a relatively simple way.

2. Review of global land cover products

In the following sections we provide brief descriptions of global products for three principal land cover types. In some cases, one product maps more than one of the categories of land cover. In these cases, we cover the multi-class product in the category to which we consider the product most relevant. Because of the importance of spatial scale and resolution discussed previously, we categorize each product by spatial resolution.

2.1. Settlements

   "All things good on this Earth flow into the City, because of the City's greatness."

Pericles of Athens, channeled by Mayor John Pappus, played by Al Pacino (City Hall, 1996)

2.1.1. The challenge of defining urban

Mapping human settlements accurately with remotely sensed imagery is challenging. In addition to the inherent difficulty of defining urban land use in terms of land cover, robust identification of land cover using remote sensing has its own challenges. Many of these challenges are related to the fact that the physical conditions and properties of land cover that influence scattering and emission of radiation vary in time, space, geography, geometry and wavelength. A second set of difficulties arises in the process of discretizing multiple scales of land cover that often vary continuously in abundance over the range of spatial scales at which sensors integrate responses into individual pixels.

The conceptual challenge of defining urban land cover is related to the lack of an agreed upon definition of urban land use. The diversity of definitions currently in use, and a discussion of the challenges of reconciling these differences, is summarized concisely by Deuskar (2015) http://blogs.worldbank.org/sustaina-blecities/what-does-urban-mean. A complementary discussion of urban and rural distinction in the context of population and mapping is given by (Dorélien et al., 2013).

In addition to the conceptual challenge of defining urban land use and human settlements, there are physical challenges related to their spatial and spectral characteristics that make the process of automated identification more difficult than may be apparent. These challenges are related to the spectral and compositional heterogeneity both within and between different settlement types worldwide. An overview of spectral heterogeneity and its importance in remote sensing of human settlements is given by (Small, 2009).

In this section of the review we discuss the most recent versions of several independent products representing human settlements on the basis of remotely sensed observations. Several of these products have evolved methodologically from earlier products not discussed here because they have been superseded by the current versions. In addition, there are other products derived from those described here, in some cases combining multiple settlement products together with other ancillary information to produce the derived product. Discussion of derived products is beyond the scope of this review, but one particular class of derived product, population maps, is discussed briefly at the end of the section as an entry point for interested readers.

2.1.2. Hectometer resolution products

The **GlobCover V2** product, produced by a research team from several European institutions coordinated by the European Space Agency (ESA) as part of the GlobCover project, contains multiple products including a 22 class thematic classification with land cover classes defined by the UN Land Cover Classification System. The 300 m resolution global land cover classification is derived from biweekly composites of MERIS optical reflectances. Both supervised and unsupervised classification algorithms were used, including spectro-temporal characterization for some land cover classes. The V2 product evolved from the earlier Global Land Cover 2000 product (Bartholome and Belward, 2005), produced by teams in affiliation with the European Commission Joint Research Centre

(JRC). In addition to land cover classes based primarily on vegetation communities, the classification contains a single class for 'Artificial surfaces and associated areas (urban areas >50%)' accounting for 0.2% of total mapped area. The validation process, described in detail by (Bicheron et al., 2008), was based on vicarious validation of 4164 point locations by a team of 15 experts using a combination of maps and high spatial resolution imagery. The stated accuracy of the classification is 58% overall with users' and producer's accuracies of 20% and 69% respectively for the urban class corresponding to human settlements. The V2 product is described in detail by (Arino et al., 2007). GlobCover products are publicly available for download at: http://due.esrin.esa.int/page_-globcover.php (access 28-08-2015).

The **MODIS 500 m Map of Global Urban Extent**, produced by a research team at the University of Wisconsin, Boston University and Princeton University, contains a single binary urban class based on a definition of 'built environment'.

> "The 'built environment' includes all non-vegetative, human-constructed elements, such as buildings, roads, runways, etc. (i. e. a mix of human-made surfaces and materials), and 'dominated' implies coverage greater than or equal to 50% of a given landscape unit (here, the pixel)." (Schneider et al., 2010).

As inputs, the 500 m resolution discrete thematic classification uses MODIS Nadir BRDF-adjusted reflectance (NBAR) (Schaaf et al., 2002) and temporal signatures of vegetation phenology derived from the Enhanced Vegetation Index (EVI) (Huete et al., 2002) in addition to a biome designation based on climate and vegetation, regional differences in urban typology and level of economic development. The classification algorithm is based on ensemble decision trees, similar to that used for the multi-class MODIS Land Cover product (Friedl et al., 2009). The decision tree classifier is trained on a combination of 1860 uniform land cover training sites corresponding to International Geosphere-Biosphere Program (IGBP) land cover classes and an independent set of 182 urban training sites. The accuracy assessment, described in detail by (Potere and Schneider 2009), was based on a comparison with analyst interpretations of a random sample of 140 Landsat-based maps of urban areas and metropolitan regions (Angel et al., 2005). The stated mean accuracy of the classification is 93% at the pixel level with $R^2 = 0.9$ at the city scale (Schneider et al., 2010). The continental average urban area of this product is 0.5% of total classified area. Additional information about product availability is available from: http://nelson.wisc.edu/sage/data-and-models/schneider.php.

More recently, the MODIS 500 m product has been refined by incorporating 250 m resolution MODIS EVI time series in addition to the 500 m MODIS NBAR reflectance. The new product, also derived from decision tree classification of spectral and temporal signatures of urban and non-urban land cover, incorporates an *a priori* urban probability surface to yield binary urban extent masks for both 2000 and 2010 (Mertes et al., 2015). Stated accuracies for the refined product in east Asia are 70% to 91% overall by country and 69% to 89% by biome. While the refined MODIS urban product is currently available only for east Asia, a global product is currently being developed (A. Schneider, Pers. Comm.).

Analyses of decadal changes in urban extent in east Asia illustrate the potential of multitemporal monitoring of human settlements. When MODIS-derived urban extent maps were compared with WorldPop disaggregated population products (described below) for 2000 and 2010, built areas were found to be expanding slower than the disaggregated populations mapped into the same areas, suggesting increased densification of urban areas in China and southeast Asia (Schneider et al., 2015). This result is particularly interesting in light of the meta-analysis of (Seto et al., 2011) which found growth rates of both population and

urban area of China and southeast Asia increasing between the 1970s and 1980s then decreasing between the 1980s and 1990s such that the overall urban land expansion rates were higher than the urban population growth. The change analysis conducted by (Mertes et al., 2015) also found growth in both "mega-agglomerations" and growth of smaller agglomerations (Schneider et al., 2015). This finding is consistent with the growth of spatial networks of lighted development throughout both China and India between 1992 and 2012 mapped by (Small and Elvidge 2013). The phenomenon of simultaneous increases in network connectivity with parallel increases in smaller settlements not connected within the larger network components has been identified as a fundamental characteristic of human settlements at regional to continental scales globally – with enormous spatial networks of large agglomerations linked by corridors with varying intensities of development (Small et al., 2011). The global consistency of scaling of spatial networks of land cover is discussed in more detail in the Global Extents section of the paper below.

Two global products based on **Stable Night Lights** have been produced by the Earth Observation Group at the National Geophysical Data Center of the National Oceanographic and Atmospheric Administration. The night lights products differ from the MODIS and GlobCover products in that they are continuous field representations of a physical property (nocturnal radiance) rather than a discrete thematic classification of land cover. The products are specifically designed to map temporally stable night light and to exclude intermittent light sources like fires. This is achieved by compositing multiple cloud-free acquisitions within time windows ranging from one month to one year. Although these are continuous fields of physical measurements, we treat them as a product because of assumptions made about temporal variability of atmospheric conditions and intermittent light sources.

The Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS) has imaged light emitted from Earth's surface since the early 1970s. A digital archive was established for DMSP-OLS data at the NOAA National Geophysical Data Center in 1992. Annual composites of temporally stable night light have been processed and distributed by the NGDC for every year since 1992. Annual composites of mean brightness provide unitless digital number (DN) values with a dynamic range of 6 bits (0–63) at a spatial resolution of 30 arc seconds (~1 km at Equator). OLS night light composites are publicly available for download from http://ngdc.noaa.gov/eog/dmsp.html (access 28-08-2015).

The Visible Infrared Imaging Radiometer Suite (VIIRS) sensor, on board the NASA-NOAA Suomi satellite, has imaged light emitted from Earth's surface since 2011. The day/night band (DNB) composite product provides calibrated radiance measurements (in units of $nW/cm^2/sr$) at a pixel resolution of 15 arc seconds (~500 m at Equator). Monthly global composites of night light brightness are produced and distributed by the Earth Observation Group at the National Geophysical Data Center since 2012. VIIRS night light composites are publicly available for download from http://ngdc.noaa.gov/eog/viirs.html (access 28-08-2015).

2.1.3. Decameter resolution products

The **Global Human Settlement Layer (GHSL)** product, developed by a team at the European Union Joint Research Centre (JRC), uses Landsat 8 multispectral and panchromatic imagery (year 2014) to distinguish built up land cover on the basis of multi-scale texture and radiometric information (Pesaresi et al., 2015) at a resolution of 38 m. An assessment of the global built-up areas in the epochs 2000, 1990, and 1975 has been completed with the same method using GLS Landsat input data.

The algorithm uses a form of supervised classification based on a symbolic machine learning approach in which several different representations of built environment are used as training sets. The

image features used as input to the classifier are multispectral signatures and textural features (Pesaresi et al., 2008) to segment potential built-up areas on the basis of spectral heterogeneity (Pesaresi and Benediktsson, 2001). The databases of built environment include OpenStreetMap, Geofabrik, Geonames, Corine Land Cover 2000 and Africover. Areas identified as built up are further subdivided into seven classes of building density on the basis of a vegetation index and a volumetric index derived from SRTM and ASTER 30 m digital elevation models. Validation of the GHSL was conducted using the Land Use/Cover Area frame Survey of the European Union (270,000 point observations circa 2012) and a collection of cartographic datasets of building footprints, primarily from Europe and the USA. Reported accuracy, based on a combination of three independent accuracy metrics, was 96% (Pesaresi et al., 2015). The same accuracy assessment applied to other remotely sensed classifications of built up areas, including GlobCover V2 and MODIS 500m, gave accuracies ranging between 95% and 97% for the European validation sites. Additional accuracy assessment using a reference set of 1505 10×10 km^2 tiles gave comparable results (Pesaresi et al., 2015). The GHSL is currently being tested at global scales and is expected to be publicly available from JRC in 2016.

The **Global Urban Footprint (GUF)** product, developed by a team at the Deutsche Zentrum für Luft- und Raumfahrt (DLR), uses X band microwave radar measurements from the TanDEM-X mission to distinguish built area from other types of land cover on the basis of backscatter intensity and its spatial texture (Esch et al., 2011). Specifically, the discovery that built environments return higher backscatter intensity values and greater variability of signal returns within windows of adjacent pixels provides an independent measure of land surface characteristics, distinct from optical properties. This allows for development of a continuous metric of 12 m resolution that can be segmented at different thresholds to produce a discrete binary classification of built up and non-built up area. Built-up area is defined as a region that shows a vertical structuring, (e.g. due to buildings or comparable infrastructure) with a certain height (Esch et al., 2013). Roads and other flat impervious surfaces are not detected. Validation, conducted using "blind photointerpretation of VHR (very high resolution) optical images by GIS experts" for the cities and surroundings of New Delhi, Buenos Aires, Munich, Nairobi and Padang, gave stated accuracies between 94.8% and 96.4% (Esch et al., 2013).

A recent comparison of the GUF and GHSL in Europe found the accuracies of these higher resolution products to be nearly double that of some lower resolution products with substantial consistencies in both urban and rural landscapes (Klotz et al., 2016). These radar-derived settlement maps have been combined with Landsat imagery to quantify the amount and form of lateral urban growth in several megacities worldwide (Taubenböck et al., 2012). Applying this type of change analysis beyond megacities to regional and continental scales offers the possibility of extending change analyses like those of (Mertes et al., 2015) to order of magnitude finer spatial scales to quantify the structure and scaling of the continent-scale spatial networks of development identified globally by Small et al. (2011). The GUF product is currently being produced at global scales and is expected to be available from DLR in 2016.

2.1.3.1. Integrated products. In addition to the aforementioned products representing human settlements and built environments on the basis of land cover, there are several products that represent ambient and resident populations as either discrete masks or continuous fields of population count or density (count/area). Despite the apparent similarity to land cover products, these integrated products are fundamentally different in the sense that they represent a phenomenon rather than a physical environment

and the phenomenon necessarily includes dimensions that are intentionally not incorporated in land cover products. Attribution of land use from land cover is similar in that it represents a phenomenon related to land cover but with additional dimensions requiring additional assumptions. Even a superficial discussion of population mapping and its relationship to land cover is beyond the scope of this review, but for reference we provide brief descriptions of some current population products that may be of interest to readers. The models are presented in a rough progression of increasing complexity of inputs, assumptions and algorithms.

GPW—Gridded Population of the World. GPW is produced by the Center for International Earth Science Information Network (CIESIN) at Columbia University as an ongoing service for the NASA SocioEconomic Data and Applications Center (SEDAC). The GPW products consist of geographic grids of population count and density produced from a collection of geographically registered census enumerations at varying levels of subnational aggregation. The gridding algorithm used to produce GPW assumes that the population within each administrative unit is distributed uniformly within the habitable land area of the administrative unit. The sole assumption of uniform distribution makes GPW an important endmember in the continuum of population models because it represents maximum dispersal and minimal clustering within the constraints of the administrative units. This is why the Lorenz curves in Fig. 1 are considered minimal estimates for the degree of spatial clustering. Because population is almost never distributed uniformly in space at kilometer scales, the true population distribution should always be more clustered than the GPW depiction. GPWv1, produced in 1997, contained 19,032 input units of nominal 1990 population. GPWv4 contains 12.5 million inputs of nominal 2010 population to yield a global grid with 30 arc second spatial resolution. Additional information, and the product, are available at: http://www.ciesin.columbia.edu/data/gpw-v4/.

GRUMP—Global Rural Urban Mapping Project. GRUMP was also produced by the Center for International Earth Science Information Network (CIESIN) at Columbia University for the NASA SocioEconomic Data and Applications Center (SEDAC). Like GPW, GRUMP products consist of geographic grids of population count and density produced from a collection of geographically registered census enumerations at varying levels of subnational aggregation. GRUMP differs from GPW in that it is an attempt to disaggregate population into rural and urban areas by estimating spatial extents of urban areas and transferring population from the surrounding unit(s) into the urban footprint and reducing the non-urban population and density accordingly. While the assumptions used for the disaggregation are more complex than the single assumption of GPW, the assumptions are fewer and simpler than those used in the Landscan and WorldPop products described below. In this sense, GRUMP is a population product of intermediate specificity and detail. A detailed description of GRUMP and its characteristics is given by (Balk 2009; Dorélien et al., 2013). The GRUMP product, and its supporting documentation, are available at: http://sedac.ciesin.columbia.edu/data/collection/grump-v1.

LandScan is an ambient population product, produced by researchers at the Oak Ridge National Laboratory of the US Department of Energy. LandScan uses a "suite of novel and dynamically adaptable" algorithms (Rose and Bright, 2014) with a variety of geospatial data and remotely sensed imagery to spatially disaggregate subnational census data on global scales at a spatial resolution of 30 arc seconds (~1 km at Equator). Like the JRC GHSL product, LandScan also uses an unsupervised classification algorithm to extract texture based metrics from remotely sensed imagery (Vijayaraj et al., 2007) but also incorporates other ancillary

information to provide estimates of ambient (as opposed to resident) population. The LandScan product has been in production, with annual updates, since 1998. Because the algorithm and inputs used to produce LandScan evolve from year to year, multitemporal change analyses using different annual releases are not recommended by the producers. Additional information about LandScan is available at: http://web.ornl.gov/sci/landscan/index.shtml.

WorldPop The WorldPop population product, produced by a team of researchers affiliated with the GeoData Institute at Southampton University, combines geospatial data, subnational census data and remotely sensed imagery in combination with a random forest disaggregation algorithm to produce a suite of spatially explicit demographic variables (Stevens et al., 2015). Early versions of the regional products (e.g. AfriPop) that have evolved into WorldPop relied primarily on global land cover products like GlobCover for disaggregation (Linard et al., 2011), but the current version incorporates a wider range of covariate inputs like VIIRS night light, digital elevation models, land cover information from Landsat, climate data and country-specific data as inputs to a Random Forest classification algorithm (Stevens et al., 2015). The resulting population and demographic parameters are gridded at ~100 m resolution. Additional information about WorldPop is available at: http://www.worldpop.org.uk.

2.1.3.2. Illustrations of settlement products. A selection of decameter, hectometer and kilometer resolution products are compared qualitatively by superposition in Fig. 6. Because settlements occupy relatively small areas at global scales, comparisons are made at regional scales using the Ganges-Brahmaputra delta for comparison with the regional illustrations of agriculture and forest products. The delta and its surrounding highlands and mountain ranges provide a representative environment for intercomparison because of the diversity of land use, the wide range of population densities, the wide range of biomes and climates. The area has many similarities to other developing tropical landscapes where much of the growth of human settlements is currently occurring. To provide a common reference for each comparison, a tri-temporal night light composite is shown for the same area at the same scale. The night light composite was produced by fusing intercalibrated OLS composites from 1992, 2002 and 2012 with a higher resolution VIIRS composite from 2012. The fusion combines the decadal change information (color) from the OLS composites with the more detailed brightness information from VIIRS (gray) to show the locations where change is observed in OLS night lights superimposed on VIIRS more detailed mapping of currently lighted areas.

At regional scale and kilometer resolution, the DMSP night light composite from 2008 is superimposed on the 2008 LandScan population density (Fig. 6a). Larger settlements appear as magenta against the green background. Areas having both high population density and high night light brightness appear pink to white. Areas of overglow where the light extends beyond the populated area (e.g. coastal cities of Chittagong and Haldia, south of Kolkata) appear magenta. Qualitatively, there is good agreement between the densely populated lowlands in LandScan and the background luminance detected by VIIRS. The utility of the less sensitive OLS night light product is apparent in the change component shown on the VIIRS composite. Despite VIIRS' higher resolution and greater sensitivity to smaller, dimmer lights, it is not available prior to 2012 so cannot provide the change information that OLS can. The increased brightness of major roads connecting larger settlements imaged by VIIRS results from both vehicle headlights and lighted bazaars along the highways where most night time activity takes place.

At regional scale and hectometer resolution, the MODIS 500 m urban product and the GlobCover urban class show similar structure for the large cities of Kolkata, Dhaka and Chittagong (Fig. 6b). In general, the MODIS product appears to resolve more fine detail within larger cities and many more small settlements in India and northern Bangladesh (red). Whereas the GlobCover product is more conservative in terms of detection of small settlements, it shows more spatial continuity within and between larger cities (green) surrounding the urban cores where both products agree (white) and where VIIRS shows the brightest lights. Neither product resolves the more dispersed rural populations like those on the coastal plain between Dhaka and Kolkata, but VIIRS shows a range of small dim lights as well as the road network connecting the larger towns. The most recent Bangladeshi census gives population densities of ~1000 people/km^2 for this area.

At local scale and decameter resolution, the JRC and DLR products show qualitatively similar extents (white) for the most brightly lighted part of Dhaka (Fig. 6c). Both products correctly exclude wetlands and agricultural areas appearing dark in VIIRS around the western periphery of the densely built up core of Dhaka. While each product resolves areas not identified by the other (magenta & green), it appears that the Landsat-derived JRC product is considerably more conservative than the DLR product. Although VIIRS resolves measurable background luminance from many of these areas, the ~500 m spatial resolution does not approach the decameter resolution of the DLR or JRC products. Because of the aforementioned tendency for buildings in rural areas to be built amongst trees, the Ganges-Brahmaputra delta provides a particularly challenging environment for settlement mapping with remote sensing.

2.2. Food production

"And the farms grew larger and the owners fewer. And the crops changed. Fruit trees took the place of grain fields, and vegetables to feed the world spread out on the bottoms: lettuce, cauliflower, artichokes, potatoes, stoop crops . . . And the owners not only did not work the farms any more, many of them had never seen the farms they owned." John Steinbeck, The Grapes of Wrath.

2.2.1. The challenge of mapping agriculture

Like mapping human settlements, mapping agriculture globally with satellite remote sensing is a daunting task. In addition to the general difficulties of all land cover identification from satellite imagery, one primary challenge arises in the definition of the term *agriculture*. Decisions must be made by the producer of any agriculture product on how to treat common situations such as multiple cropping, intermittently fallow land, and multiple land uses. Another set of challenges arises from the global diversity of agriculture and the concomitant range of spatial, spectral and temporal patterns it can assume. Compounding the complexity of agricultural patterns is the issue of non-uniqueness: some of these patterns are spectrally and temporally indistinguishable from the background land cover.

As for the other portions of this study, the scope of the agriculture portion of this review is limited to attempts to estimate total global extent. For this reason, we are forced to avoid discussion of more specific agricultural remote sensing work. For example, global efforts to map irrigation (e.g. (Biradar et al., 2009; Thenkabail et al., 2009) (Siebert et al., 2005; Salmon et al., 2015)), field size (e.g. White and Roy 2015; Fritz et al., 2015), yield (e.g. Iizumi and Yokozawa, 2014), cropping intensity (e.g. Ray and Foley 2014; de Jong et al., 2013), long term cropping trends (e.g. Cook and Pau, 2013; Fensholt et al., 2012), abandonment (Alcantara et al.,
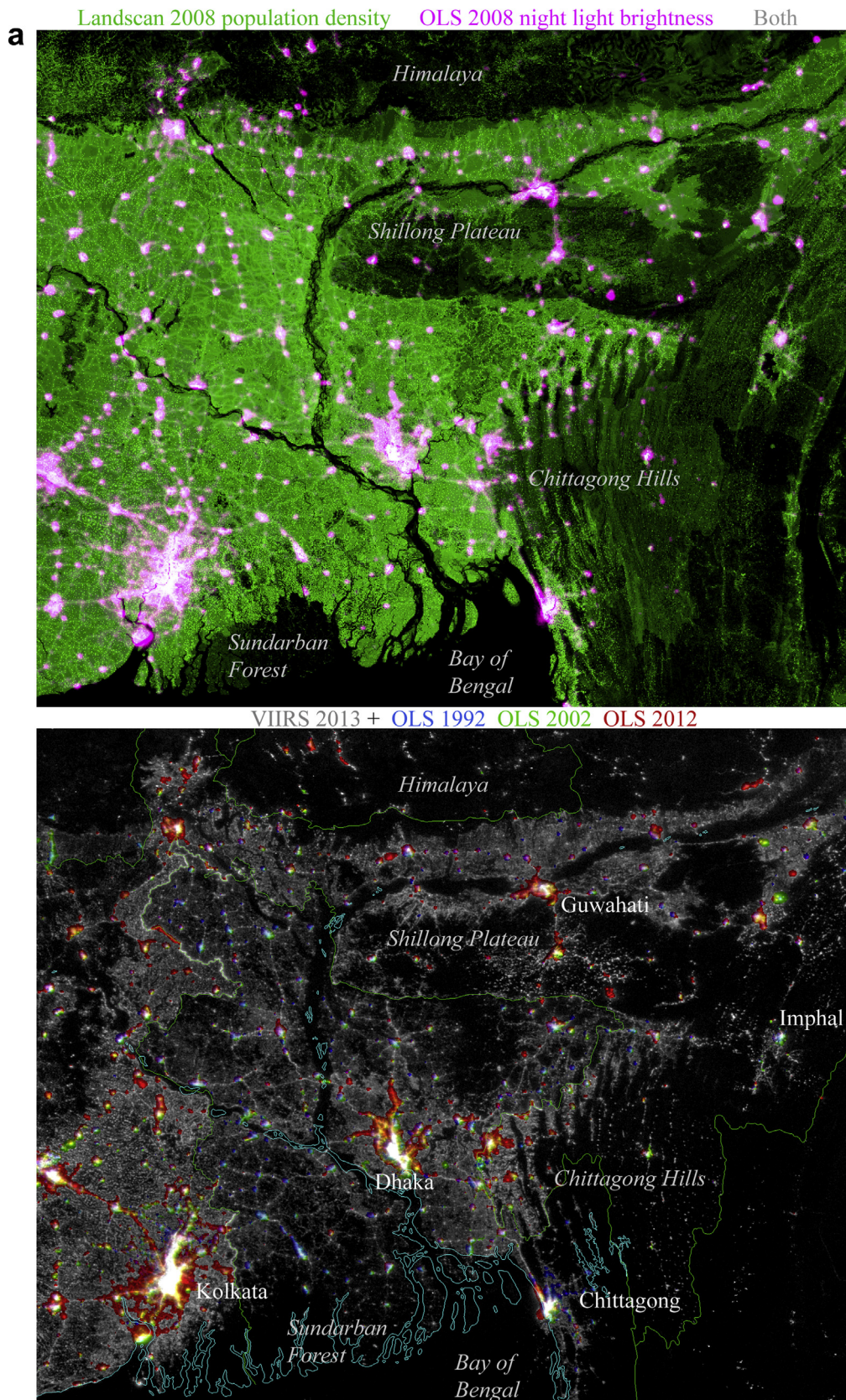
Fig. 6. (a) Kilometer product comparison with night light change for the Ganges Brahmaputra delta and surroundings. Both density and brightness vary continuously on Log$_{10}$ scale. Enlarge to see full resolution. (b) Hectometer product comparison with night light change for the Ganges-Brahmaputra delta. Enlarge to see full resolution. (c) Decameter product comparison with night light change image for Dhaka, Bangladesh. The JRC product is a continuous probability of human settlement and the DLR product is a discrete mask of built up areas. Cyan box shows location of Dhaka image in Fig. 2. Enlarge to see full resolution.
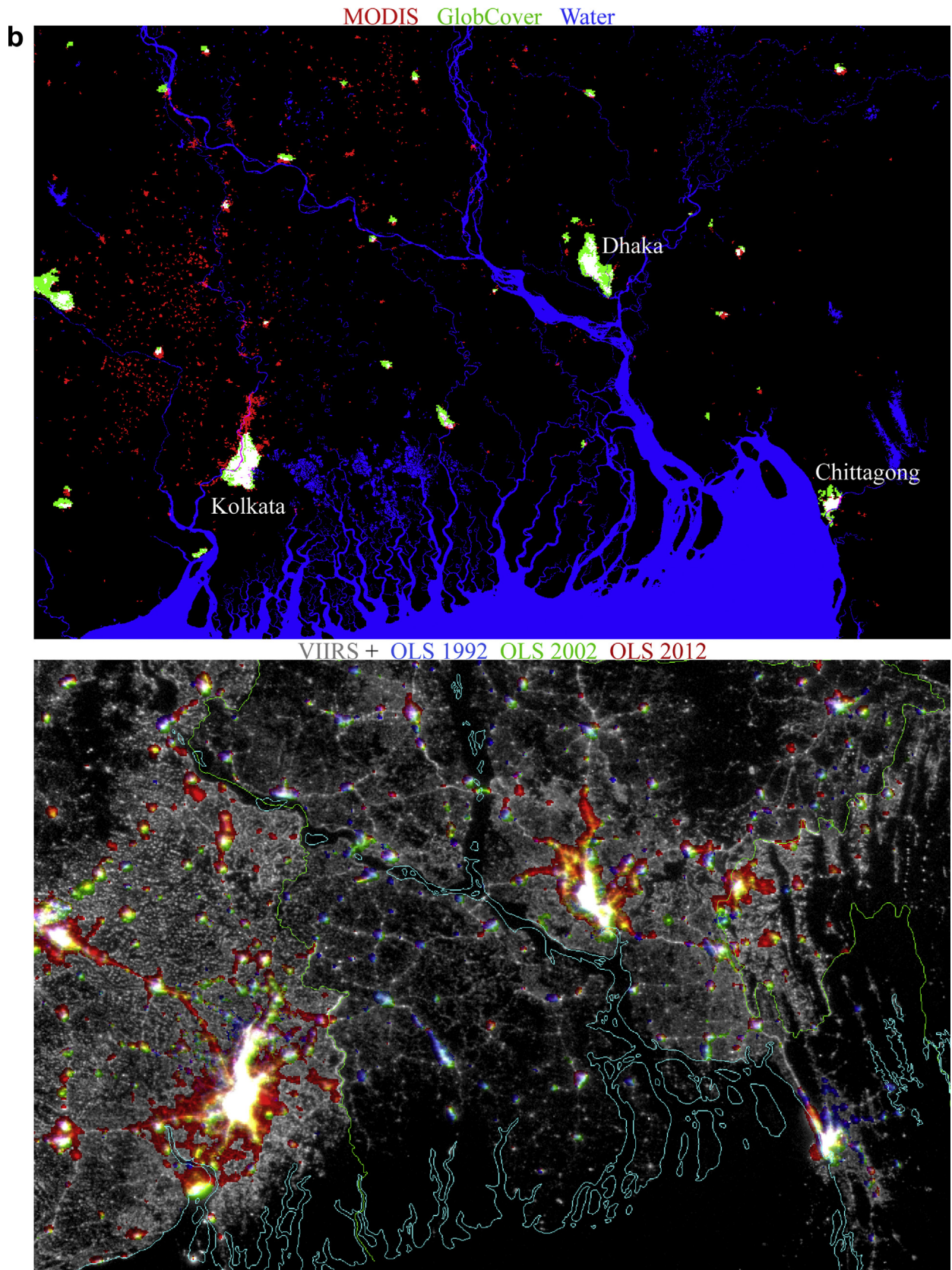
MODIS   GlobCover   Water

VIIRS +   OLS 1992   OLS 2002   OLS 2012

Fig. 6. (Continued)

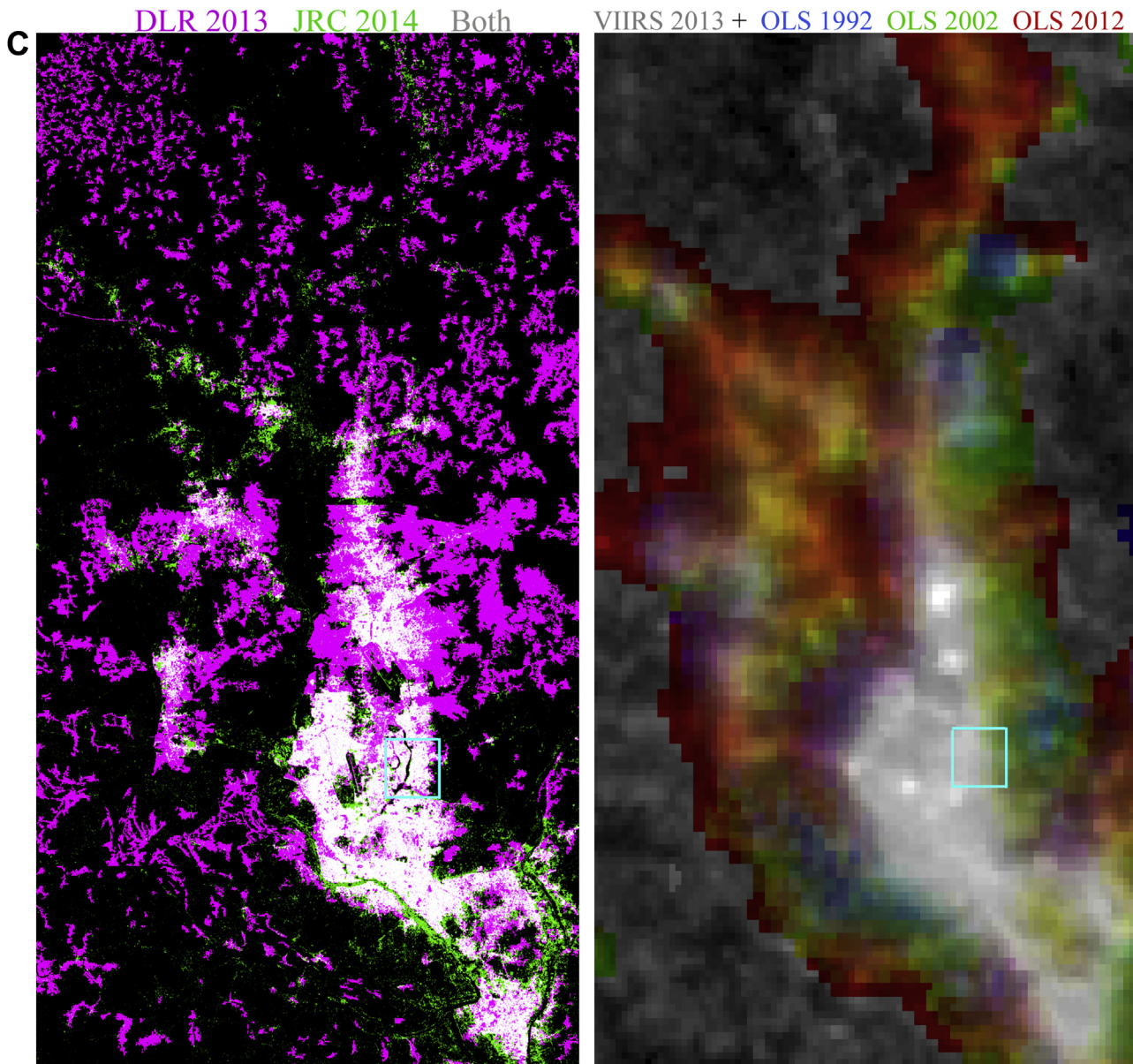**C**　　DLR 2013　JRC 2014　Both　　　VIIRS 2013 +　OLS 1992　OLS 2002　OLS 2012

Fig. 6. (Continued)

2012), land degradation (de Jong et al., 2011), or distribution of individual crops (e.g. Monfreda, 2008; Leff et al., 2004) will not be discussed. We also exclude cropland datasets created as one layer in full land cover classification schemes (e.g. Loveland et al., 2000; Hansen et al., 2000; Friedl et al., 2002; Bartholome and Belward, 2005; Arino et al., 2007), although some of these products are incorporated into the agriculture-specific products discussed here..

2.2.2. Agricultural survey data

The following two datasets are not based on remote sensing. However, all of the agricultural remote sensing products described here rely on at least one of these surveys for calibration and/or validation. The survey data are presented to give context to the remote sensing products which follow.

The **Food and Agriculture Organization of the United Nations** (FAO) collects and archives country-level statistics on a wide range of agricultural variables. These data are obtained using government surveys and cross-checked by FAO experts. The most relevant statistic for intercomparison with remote sensing studies is the global arable land area, defined by the FAO as

'land under temporary agricultural crops (multiple-cropped areas are counted only once), temporary meadows for mowing or pasture, land under market and kitchen gardens and land temporarily fallow (less than five years). The abandoned land resulting from shifting cultivation is not included in this category. Data for "arable land" are not meant to indicate the amount of land that is potentially cultivable' (FAO).

Currently, the most recent available estimate for global arable land is 15.6 million km² for the year 2012. Agricultural survey data have been shown to exhibit a number of biases and inaccuracies, particularly in developing countries (e.g. Carfagna et al., 2013; Carletto et al., 2013, World Bank, 2010, FAO et al., 2012), but in many cases still represent the best available calibration for remote

sensing studies. FAO country level arable land estimates are used on some level to calibrate total agricultural area for all of the following remote sensing products except the Pittman et al. MODIS global cropland estimates. FAO data are publicly available online at: http://www.faostat3.fao.org/ (access 28-08-2015).

The **United States Foreign Agricultural Service (FAS) Production Supply & Distribution (PSD) Database** provides country-level crop area estimates. These estimates are compiled by several agencies of the U.S. federal government and include data from agricultural attaches stationed at U.S. embassies worldwide, commodity analysis, official country statistics, international organizations, and some remote sensing data. The PSD database is produced on the basis of marketing year (whereas the FAO database is produced on the basis of calendar year). Because of this, some double cropped land is counted twice in area totals. The PSD database also reports the European Union as one entity. The Pittman et al. MODIS Global Cropland product calibrates its estimates to the FAS PSD database. FAS PSD data are publicly available online at: https://apps.fas.usda.gov/psdonline/ (access 28-08-2015).

2.2.3. Kilometer resolution products

The **EarthStat Global Cropland 2000** product (Ramankutty et al., 2000), produced by a research team at the University of Minnesota and the University of British Columbia, contains estimates of sub-pixel proportions for cropland and pasture area. Inputs to the 10 km resolution global continuous field product include discrete classifications from the MODIS global land cover product (Friedl et al., 2002) and GLC2000 (Bartholome and Belward, 2005), as well as a compilation of agricultural census data from 15,990 administrative units. The remote sensing classifications were used to spatially disaggregate the agricultural area within each administrative unit. Confidence intervals were estimated by bootstrapping. The Earthstat estimate of total global area of croplands was 15.0 million km^2 and the global area of pastures was 31.5 million km^2 in the year 2000. Validation was

performed by comparison to existing regional products and to FAO estimates for total global cropland and pasture areas. Subsequent studies such as (Foley et al., 2011) have used the Earthstat product as the basis for further analysis on crop-specific areas, yield gaps, and fertilizer use. The Earthstat dataset is publicly available for download at: http://www.earthstat.org/ (access 28-08-2015).

The **History Database of the Global Environment (HYDE) version 3.1** (Klein Goldewijk et al., 2010), produced by a research team at the Netherlands Environmental Assessment Agency, contains historical maps of sub-pixel proportions of cropland and pasture for the last 12,000 years. The 10 km resolution continuous field product disaggregates 1990–2000 country level FAO cropland and pasture area statistics using land cover inputs from the DISCover version 2 dataset and the GLC2000 discrete classification (Klein Goldewijk et al., 2006). Using current satellite-based estimates as the baseline, HYDE 3.1 attempts to hindcast changes in cropland extent from 2000 CE back to 10,000 BCE. Several assumptions about geographic favorability for agriculture based on population density, soil quality, distance to water, terrain slope, and temperature are used in the HYDE model and given increasing weight as model time diverges from the satellite era. The HYDE 3.1 model estimate for global cropland area in 2010 is 15.3 million square km, matching the FAO estimate. Uncertainties of the HYDE 3.1 model are presented in detail in (Klein Goldewijk and Verburg, 2013). HYDE 3.1 is the only model discussed here which attempts to estimate global agriculture before the satellite era. The HYDE 3.1 dataset is publicly available for download at: http://www.themasites.pbl.nl/tridion/en/themasites/hyde/ (access 28-08-2015).

The **IIASA-IFPRI Cropland Map** (Fritz et al., 2015), produced by a research team from 32 institutions coordinated by the International Institute for Applied Systems Analysis (IIASA) and the International Food Policy Research Institute (IFPRI), contains a single map of cropland likelihood for the year 2005. The 1 km resolution global cropland probability product was produced by
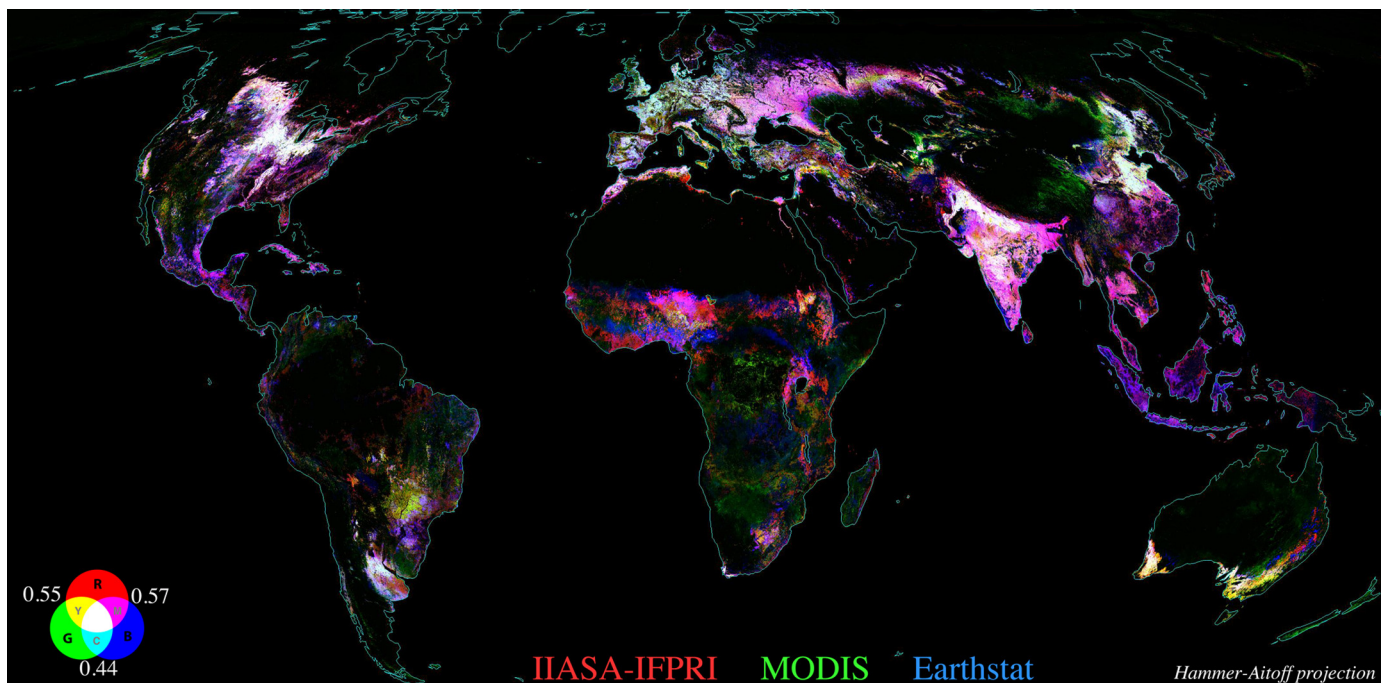

Fig. 7. (a) Global crop probability comparison. Brightness of each RGB channel is proportional to crop probability for the corresponding product, with maximum saturation at 75%. Brighter primary colors (RGB) indicate higher probability in one product. Brighter secondary colors (CMY) indicate agreement between two products. Shades of gray and white indicate comparable probabilities of all three products. Pairwise correlations shown on color wheel. (b) Crop probability maps for the Ganges-Brahmaputra delta. Additive color scheme similar to (a) but with different channel assignments. All three products correctly exclude large cities, river channels and forested uplands but differ in percentage estimates. Correlations not given because GLC is binary.
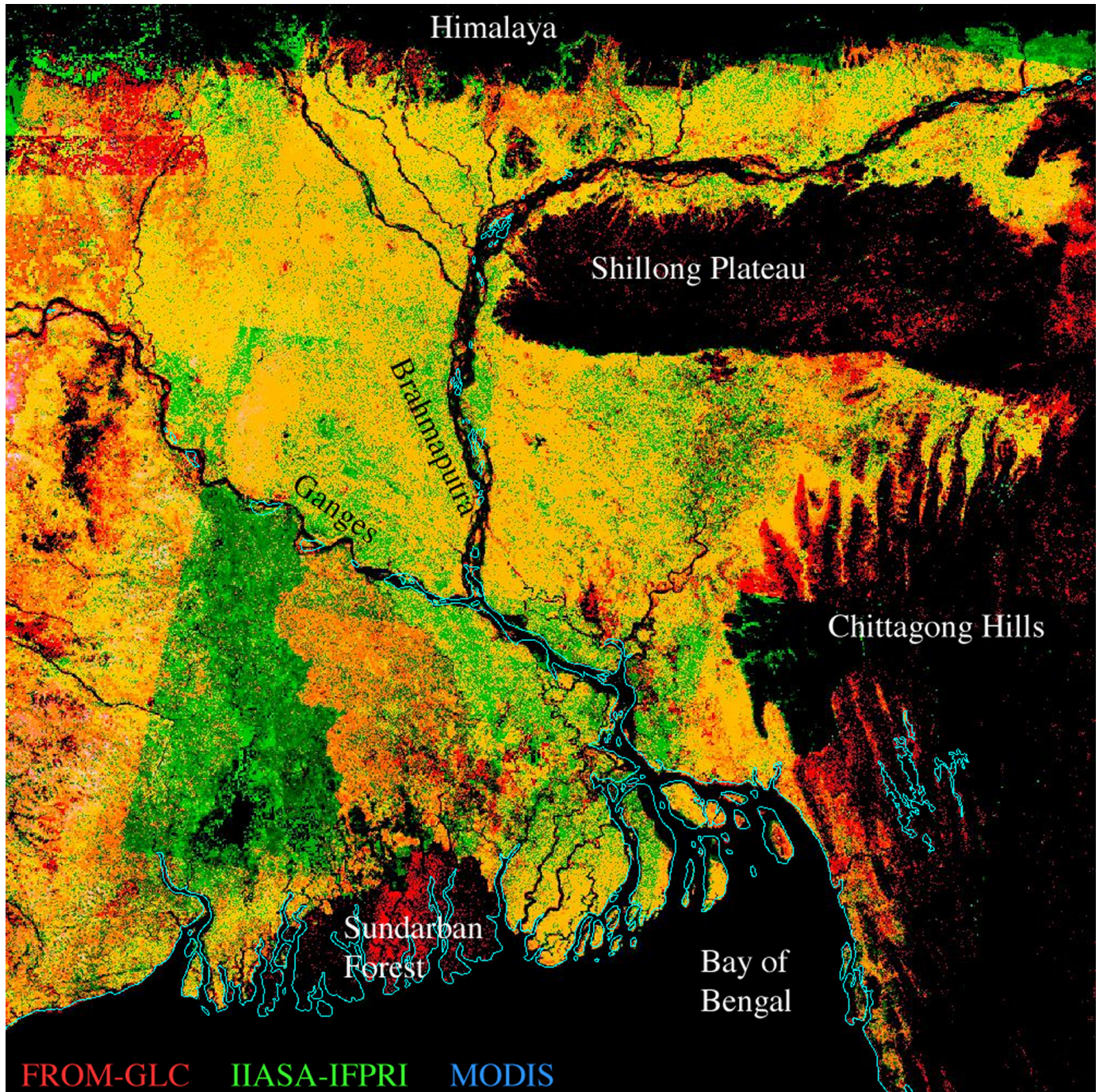
Fig. 7. (Continued)

synthesizing 5 global, 3 regional, and 14 national cropland masks, including several of the global datasets described here. The IIASA-IFPRI Cropland Map was calibrated to FAO national statistics and IFPRI subnational cropland data (You et al., 2014) averaged over the years 2003–2006. Detailed comparison and consistency assessments against the EarthStat cropland product were performed. The datasets were validated by crowd-sourcing against high resolution Google Earth imagery, in which volunteer participants "were asked to choose the dominant land cover type from among ten sample land cover classes including one class for cropland and one for a mosaic of cropland and natural vegetation". 10,626 cropland percentage pixels were used in the final validation data set. Expert opinion was used to resolve areas of disagreement. Stated accuracy of the IIASA-IFPRI Cropland Map is 82.4%. No global estimate for total cropland area is included with the analysis, but calculation from summing the number of cropland pixels in an equal area projection of the global 1 km per-pixel cropland percentage map available online yields 16.3 million square km. The FAO estimate for global arable land in 2005 is 15.3 million square km. The IIASA-IFPRI Cropland map is currently being used for monthly reporting

and crop monitoring products by the Group on Earth Observations (GEO) Global Agricultural Monitoring Initiative (GEOGLAM). The IIASA-IFPRI Cropland Map is publicly available for download at: http://www.geo-wiki.org/ (access 28-08-2015).

2.2.4. Hectometer resolution products

The **MODIS Global Cropland** products (Pittman et al., 2010), produced by a research team from South Dakota State University and the University of Maryland, contains 250 m sub-pixel cropland probability and 250 m discrete cropland/not-cropland layers for 222 tiles covering Earth's land surface. 1 km resolution global mosaics of the probability and cropland/not-cropland datasets, aggregated from the 250 m products, are also available. The 250 m cropland probability layer was based on repeated supervised classification trees of 9-year time series of 39 different metrics, all derived from MODIS data. The training data used was primarily the Landsat Geocover data set, a global discrete land cover classification (Tucker et al., 2004). Several other regional discrete land cover or binary crop/not-crop datasets were also used. After cropland probability was calculated based on degree of convergence of the repeated classifications, the binary cropland/not-cropland maps were derived by choosing a probability threshold to match the country level median 2000–2008 harvested area for all field crops from the FAS PSD database. Validation was performed by comparison to 5 satellite-derived global land cover products. The MODIS Global Cropland product was incorporated in the data stream used by FAS analysis to assess crop conditions. The MODIS Global Cropland datasets is publicly available for download at: http://www.glad.geog.umd.edu/projects/croplands/ (access 2008-08-2015).

2.2.5. Decameter resolution products

The **Finer Resolution Observation and Monitoring, Global Cropland** (FROM-GC, Yu et al., 2013) contains a 10 class thematic classification optimized for cropland accuracy. The 30 m resolution global land cover classification is created through merging results from five datasets: two versions of the associated 30 m global land cover product (FROM-GLC, Gong et al., 2013), the 250 m cropland probability map developed by (Pittman et al., 2010), a global land cover validation sample database from (Zhao et al., 2014), and FAO country level agricultural statistics. A support vector machine-based discrete classification of over 8900 Landsat scenes was used. The Zhao et al. (2014) validation dataset is a collection of 38,664 samples produced by visual interpretation of Landsat images, MODIS EVI time series, and meter resolution imagery from Google Earth. The 250 m cropland probability mask and FAO statistics are described above. FROM-GC estimates a global cropland area of 15.3 million square km for the year 2010. The FAO estimate for global arable land in 2010 is also 15.3 million square km. The FROM-GC dataset is publicly available for download at: http://data.ess.tsinghua.edu.cn/ (access 2008-08-2015).

Fig. 7a shows global maps of three cropland products. Large, concentrated, highly productive areas such as the North American corn and wheat belts, the Great Central Valley of California, the Argentinean Pampas, the North China Plain, the Indo-Gangetic Plain, and southwest Australia show widespread agreement. More dispersed agricultural regions and areas with less robust survey data present more of a challenge. Fig. 7b shows a regional comparison for three products in the Ganges-Brahmaputra delta. Note that two of the products are the same as in Fig. 7a, but the colors to which they correspond have been changed for clarity. All three models successfully separate the intensively cultivated alluvium surrounding the river channels from the forested

highlands of the Shillong Plateau and the Chittagong Hills. Spatial distributions of sub-pixel cropland fractions are more variable.

2.3. Forest cover

> "Whose woods these are I think I know. His house is in the village though". Robert Frost, Stopping by Woods on a Snowy Evening.

2.3.1. Forest as land cover

As in the cases of urban and agricultural areas, forest cover mapping on the global scale has several intrinsic challenges. The nature of these challenges are similar to those already discussed: uncertainty as to the definition of the term "forest", diversity in canopy structure and phenology, difficulty distinguishing between forest cover and non-forest vegetation (non-uniqueness), and year-to-year forest regrowth and clearing. Additional forest-specific challenges include forest degradation within static forest boundaries and difficulty distinguishing between natural forest and silviculture. The general land cover challenges related to scale also apply. The challenges of forest cover mapping, and the importance of increasing spatial, spectral and temporal resolution, are especially apparent when comparing early development of monitoring strategies (e.g. Skole and Tucker 1993) and early syntheses (Skole et al., 2004) with current capabilities and observations described here.

The scope of the forest cover portion of this review is limited to attempts to specifically map total global forest extent. For this reason, we are forced to avoid discussion of more specific forest cover studies, for example regional products (e.g. Verhegghen et al., 2012; Kempeneers et al., 2012; Schepaschenko et al., 2011; Potapov et al., 2011) or products specific to forest type (e.g. Giri et al., 2007; Hamilton and Casey, 2014). We also exclude forest cover layers produced as part of complete land cover classifications (e.g. Bartholome and Belward, 2005; Tateishi et al., 2011; Friedl et al., 2010) and methodologies designed for forest cover monitoring (e.g. Asner et al., 2009). We display and briefly describe only the currently available forest cover products. Detailed analyses of differences between global forest cover products exist in the literature (e.g. Schepaschenko et al., 2015; Seebach et al., 2012).

2.3.1.1. Kilometer resolution products. The **IIASA Hybrid Global Forest Product** (Schepaschenko et al., 2015), developed by a research team at 11 institutions, provides three global continuous estimates of forest probability at 1 km for the year 2000. Two of these are produced to be consistent with FAO statistics and one is a "best guess" estimate which is not consistent with the FAO. The 1 km resolution forest maps are made using geographically weighted regression on eight different global forest products and 7 regional forest products. The stated accuracies of the products are 93% for the hybrid "best-guess" map, 85% for the hybrid map calibrated by country and 86% for the hybrid map calibrated by FAO region. A crowdsourcing method similar to that of (Fritz et al., 2015) was used for validation. The IIASA Hybrid Global Forest Product dataset is publicly available at: http://biomass.geo-wiki.org/ (access 28-08-2015).

2.3.2. Hectometer resolution products

The **MODIS Vegetation Continuous Fields version 5** product (MODIS VCF, DiMiceli et al., 2011) developed by a research team at the University of Maryland, will contain sub-pixel area estimates for tree cover, herbaceous cover, and bare cover. Currently, only the percent tree cover component of the product is available. The 250 m MODIS VCF product is generated from a regression tree

based upon the full 7-band MODIS reflectance data. Theoretical basis for the algorithm is described in (Hansen et al., 2002). The MODIS VCF tree cover layer is publicly available at: http://landcover.org/data/vcf/ (access 28-08-2015).

2.3.3. Decameter resolution products

The **Global Forest Change 2000–2013** product (Hansen et al., 2013), developed by a research team at 6 institutions, contains global estimates of: tree canopy cover for the year 2000, forest cover loss 2000–2013, forest cover gain 2000–2012, year of gross forest cover loss event, and circa 2000 and 2013 cloud free Landsat image composites. The 30 m dataset is based on time-series analysis of over 650,000 Landsat scenes based on a set of bagged decision trees. The dataset is publicly available online at: http://earthenginepartners.appspot.com/science-2013-global-forest/download_v1.1.html (access 28-08-2015). Validation was based on stratified random sampling of 120 m blocks in each forest biome. 1500 total blocks were interpreted by experts and used in the final validation dataset. The overall accuracy of global forest loss and gain maps is stated to be 99.6% and 99.7%, respectively. Landsat-based forest cover change has also been extended retrospectively to 1990 by (Kim et al., 2014). The Global Forest Change 2000–2013 product finds 2.3 million km^2 of forest lost due to disturbance and 0.8 million km^2 new forest established, with 0.2 million km^2 experiencing both loss and subsequent gain during the study period. Of the four climate domains considered by the study (tropical, subtropical, temperate, boreal), the greatest forest loss and gain was within the tropical domain. A total of 32% of global forest cover loss occurred within tropical rainforest ecozones. This dataset is expected to be available at: http://landcover.org/data/landsatFCC/ (access 28-08-2015).

The **ALOS/PALSAR Global Forest/Non-Forest Map** (Shimada et al., 2014), produced by a research team coordinated by the Japan Aerospace Exploration Agency (JAXA), contains discrete global forest/non-forest maps for the years 2007–2010 at 25 m resolution. Also available from the same group, the FNF 100 product aggregates the results of the 25 m classification to 100 m resolution. Pixels are binned into 4 categories (0–25%, 26–50%, 61–75%, 76–100% forest cover) based on the number of 25 m forest pixels inside each 100 m pixel. The ALOS/PALSAR product is derived exclusively from L-band radar backscatter images, as opposed to the optical imagery used in the other products described here. The SAR data were collected by the JAXA Advanced Land Observing Satellite (ALOS) Phased Array type L-band SAR (PALSAR) sensor over the years 2007–2012. Forests can be distinguished from other land cover types in L-band SAR imagery on the basis of differences in radar reflectivity and geometric arrangement of tree trunks and branches. Validation was done with 1456 forest and 2548 non-forest validation points assembled from the Degree Confluence Project (DCP) and Google Earth Imagery (GEI). Stated accuracy varied between 90 and 92% for GEI validation points and 82–88% for DCP validation points, depending on the year. The ALOS/PALSAR Global Forest/Non-Forest Map finds a decrease in forest cover of 31.9 million ha from 2007 to 2008 and an overall decrease over the years 2007–2010 of 1.6 million ha. The 25 m and 100 m datasets are publicly available at: http://www.eorc.jaxa.jp/ALOS/en/palsar_fnf/fnf_index.htm (access 28-08-2015).

The **Landsat Vegetation Continuous Fields** (Landsat VCF) product (Sexton et al., 2013), produced by a research team at the University of Maryland, will include a global 30 m map of sub-pixel area covered by woody vegetation greater than 30 m in height. The data will include two nominal epochs, 2000 and 2005, based on the Global Land Survey (GLS) collection of Landsat data. The Landsat VCF product rescales the MODIS VCF Tree Cover dataset using a regression tree algorithm based on all 7 bands of Landsat TM/ETM+ data. Validation was performed using 4 small-footprint LiDAR

measurements. Root Mean Square Error (RMSE) of Landsat VCF relative to LiDAR was stated to be 17%. The Landsat VCF product is expected to be available at: http://landcover.org/data/landsatTree-cover/ (access 28-08-2015).

Fig. 8 shows the global distribution of forests from the MODIS Vegetation Continuous Fields product (top) and a regional comparison of three forest products for the Ganges-Brahmaputra delta and surrounding uplands. (bottom). The global map successfully identifies major global tropical forest areas in the Yucatan Peninsula, Amazon Basin, Congo Basin, and southeast Asian highlands as well as boreal forests in North America and Eurasia. In the regional comparison, all three products successfully identify the major forested areas of the Shillong Plateau, Chittagong Hills, and Sundarbans mangroves – although some variability between fractional estimates of the three products is also apparent.

3. Global extent comparisons

The land cover products described here provide complementary, spatially explicit depictions of the global extent of human settlements, agriculture and forest cover. In addition to showing the geographic distributions, they allow comparative analyses of the total area of each land cover type. However, because of varying definitions, detection thresholds and accuracy limitations, a simple comparison of total areas could be misleading. Given the resolution limits and scale issues discussed in the Background section, a comparison of *size distributions* may be more informative. Larger areas of spatially contiguous land cover can generally be mapped more accurately than smaller interspersed patches – however, smaller interspersed patches account for a considerable fraction of each land cover category so they cannot be ignored. In addition, the relative distributions of different size patches can inform our understanding of the land use processes responsible for the distribution of each type of land cover. For these reasons, we compare the size distributions of spatially contiguous patches of several of the land cover products discussed here. The purpose of the comparison is to illustrate similarities and differences in the overall distributions – rather than specific geographic disparities. A rigorous comparison of all the products discussed here is beyond the scope of this review. For practical reasons, we have chosen some of the kilometer resolution products for this initial comparison. Before presenting the extent comparisons, we provide some background on rank-size distributions and their utility for quantifying spatial distributions and their scaling properties. Scaling properties of land cover and other dynamic systems provide important constraints for testing hypotheses about their structure and evolution, even when their spatial structure becomes arbitrarily complex.

3.1. Rank-Size distributions

The global products discussed here contain large numbers of mapped units and cover very large areas relative to their spatial resolutions. As a result, the distributions of the mapped units can span several orders of magnitude in size and number. The rank-size distribution provides a convenient way to represent collections of large numbers of objects spanning wide ranges of sizes. When objects are sorted by size and plotted by ordinal rank, and both rank and size span multiple orders of magnitude, the resulting distribution is often found to be heavy tailed, with increasing numbers of smaller objects and decreasing numbers of larger objects. In dynamic systems characterized by growth and decay processes, the properties of these distributions can inform understanding of the processes responsible for producing them (Sornette, 2003).
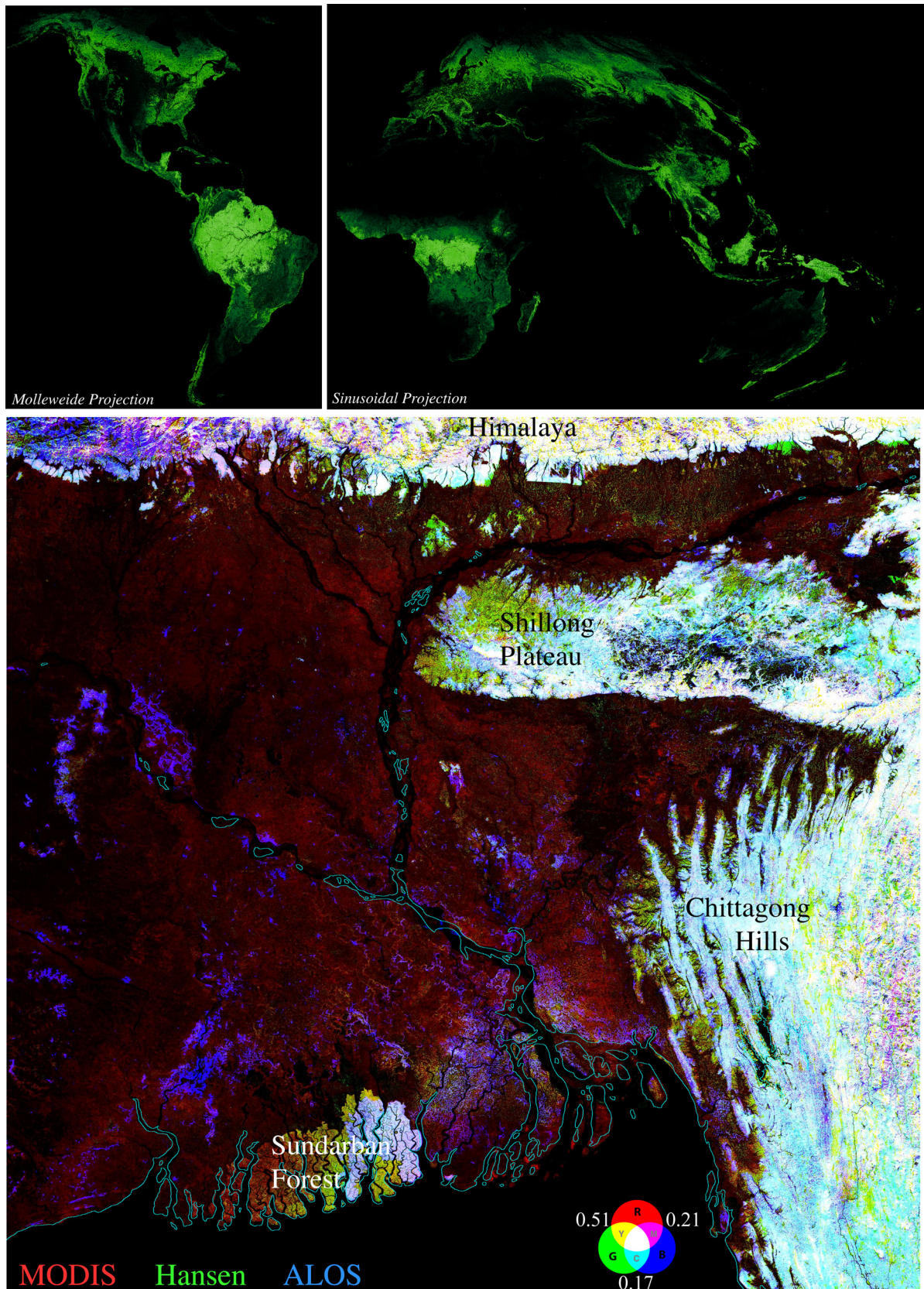
Fig. 8. Global tree cover distribution and regional forest product comparison. Fractional tree cover from Vegetation Continuous Fields (top) shows increasing densities of tree cover as lighter green. Forest product comparison on the Ganges-Brahmaputra delta (bottom) generally identifies forested uplands correctly but with differing percentage of cover. Lateral variation in Sundarban mangrove forest density results from eastward increase in canopy closure due to salinity gradient from Ganges-Brahmaputra discharge. Additive colors and pairwise correlations as in Fig. 7a.

Many of the land cover products we consider are given as continuous fields so they must be discretized for comparison with traditional maps in which land cover is represented by discrete spatial elements of varying size and shape. Within continuous fields, spatially contiguous patches can be quantified by the process of segmentation. Each patch, also called a segment, is defined on the basis of spatial connectivity of pixels. In the case of discrete thematic land cover classifications, each class is segmented into spatially contiguous patches by the classification process itself. In the case of continuous fields, a threshold must be imposed to create a binary segmentation into patches of pixels with values above and below the threshold. The choice of threshold determines the decision boundary and therefore the distribution of segment or patch sizes that results. In most cases, the optimal choice of threshold is not obvious. Rather than choosing a single threshold based on arbitrary or ad hoc criteria, a set of successive thresholds can be applied to the continuous field to determine the sensitivity of the resulting discrete maps to the threshold chosen. The effect of different thresholds is reflected in changes in the resulting segment size distributions.

For each threshold, the rank-size distribution of segments can span several orders of magnitude in area and number. These types of heavy tailed distributions are often represented with power laws of the form

$$s_r = k \ r^{-\alpha}$$

where the size, s_r, of each object in a sorted set is given as a function of its ordinal rank, r, an exponent, $-\alpha$, and a constant, k. The exponent represents the slope of the rank-size distribution when the Log_{10} of the size of each object is plotted against the Log_{10} of its ordinal rank (largest to smallest). The physical meaning of the exponent, or slope of the distribution, is related to total area of objects of different sizes. Larger negative exponents imply greater slopes corresponding to distributions in which larger objects account for more total area than smaller objects. Smaller negative exponents correspond to the opposite case where smaller objects account for more total area than larger objects. A distribution with an exponent of -1 corresponds to the transitional case of a uniform size distribution in which objects in each size range account for the same total area. In the case of segment areas, the slope of the rank-size distribution, given by the exponent of the best-fit power law, indicates whether larger segments account for more or less total area than smaller segments.

The rank-size distribution can provide a relatively simple and informative description of the structure of arbitrarily complex spatial patterns. In the case where the distribution has a consistent form, with segments spanning a wide range of sizes, the shape of

the distribution can inform our understanding of the process(es) that may be responsible for the formation of the spatial patterns. By combining the process of successive thresholding with estimation of the slope of the resulting rank-size distributions, it is possible to quantify the spatial structure of a continuous field as a function of different thresholds (Small et al., 2011). The resulting distributions, and their sensitivity to threshold, provide a fundamental description of the scaling properties of the continuous field – analogous to the scaling properties of classes in a discrete classification. The evolution of the size distribution of a population of objects (e.g. settlements) informs our understanding of competing processes responsible for the growth and evolution of the population.

We use rank-size distributions of several of the land cover products described here to compare their global extents and scaling properties. For continuous fields, we impose multiple thresholds and compare the resulting rank-size distributions. For each rank-size distribution of segment sizes, a power law was fit using the maximum likelihood method of (Clauset et al., 2009). This procedure repeatedly computes maximum likelihood estimates of the power law exponent for increasing subsets of the upper tail of the distribution and identifies a lower tail cutoff coinciding with the minimum of the Kolmogorov-Smirnov goodness-of-fit statistic. This cutoff gives an indication of the domain (range of spatial scales) over which the rank size distribution can be plausibly described with a power law. A lower tail cutoff is used because power laws are often defined only for the upper tail of a distribution when detection limits make the lower tail unreliable. Examples of rank-size distributions and the best fit exponents and lower tail cutoffs of the corresponding power laws are shown for settlements, agricultural areas and forests in Figs. 9, 10 and 11 respectively. Despite the difference in thresholds, the shapes of the resulting rank-size distributions are very similar with exponents close to -1 for the best fit power law of each. In each case, the most obvious effect of increasing the threshold is to reduce both the size and number of segments. However, the slopes do not appear to be sensitive to the thresholds over most of the range of each product. For settlements defined by multiple thresholds of both night light brightness and ambient population density, this result holds for each of the three longitudinal sectors (Americas, Asia and Europe/ Africa) and for the combined global distributions. For forests, defined by varying fractions of tree cover, the result holds at continent scales and for the combined global distributions. For agricultural areas, defined by four global products, the result holds for two different fraction thresholds, although more
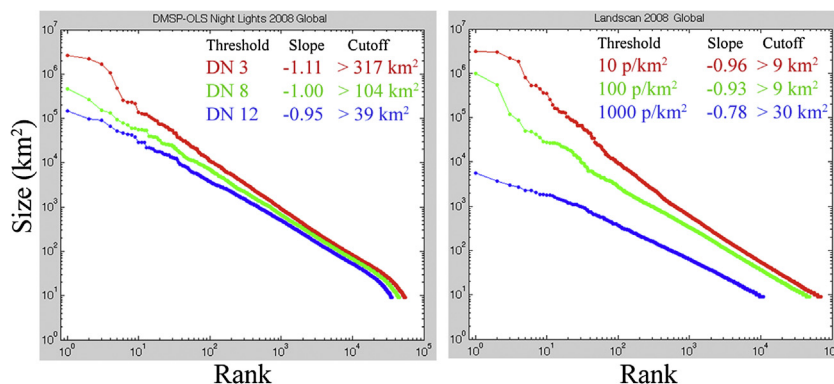

Fig. 9. Rank-size distributions for stable night lights and population density in 2008. Increasing brightness and density thresholds reduces size and connectivity of the spatial networks and reduces the slopes (power law exponents) as increasingly smaller patches and networks account for more of the total area. Night lights have larger lower bound areas because detection limits of small, dim lights causes distributions to roll off in the lower tails.
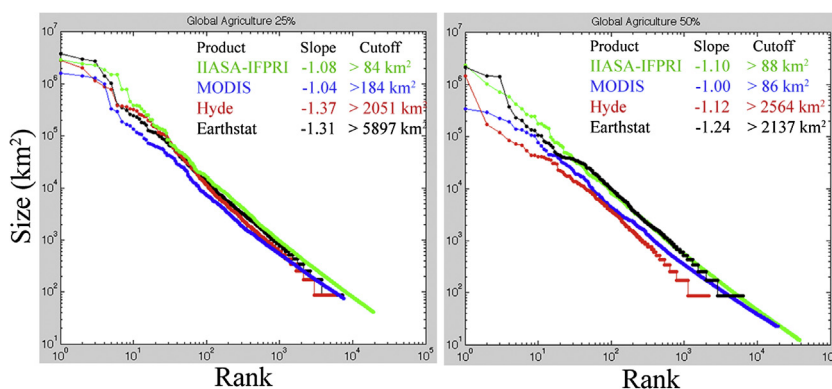
Fig. 10. Global rank-size distributions for agricultural areas. The two lower resolution products (Hyde & Earthstat) have greater slopes and larger lower bounds suggesting fundamentally different distributions than the MODIS and IIASA-IFPRI products. For both 25% and 50% agriculture thresholds the greater slopes and larger lower bounds of the lower resolution models suggest that smaller agricultural areas may be under-represented.

consistently for the two higher resolution products. The specifics and implications are discussed below.

3.2. Settlements

City size distributions, defined on the basis of population within administrative boundaries, are often described with power laws. Auerbach (Auerbach, 1913) made the initial observation that the product of a city population and ordinal rank is approximately constant. Lotka (Lotka, 1941) later observed a hyperbolic rank-size relationship for U.S. city populations, noting that the slope of the Log_{10} rank-size plot was not exactly −1 but −0.93 with some of the larger cities being smaller than predicted. Zipf later estimated that the exponent of the power law is also close to −1 for U.S. cities (Zipf, 1942), a finding that also holds for the frequency of usage of words, sizes of firms, and a variety of other socioeconomic observables (Zipf, 1949). The special case of a power law distribution with an exponent of −1 is often referred to as Zipf's Law. The assertion of a universal power law for city size is controversial because the linearity and slope of the power law rank-size distribution appears to vary through time and among countries (Gabaix et al., 2004; Ioannides and Overman, 2003; Nitsch, 2005; Overman and Ioannides, 2001; Pumain, 2004; Rosen and Resnick, 1980; Soo, 2005).

Rank-size distributions are compared for DMSP-OLS night light brightness and for LandScan population density, circa 2008, using the progressive segmentation method (Small et al., 2011), and power laws are fit using the method of (Clauset et al., 2009). The distributions, shown in Fig. 9, diminish in size and area, as would be expected when higher thresholds cause segments to shrink and disappear. The slope of each distribution also diminishes somewhat with increasing threshold, indicating that the fragmentation of larger segments in the upper tail has a more pronounced effect than the attenuation of smaller segments in the lower tail. All three of the brightness thresholds, and two of the three density thresholds, yield statistically significant fits to the power law distribution, spanning at least four orders of magnitude in segment size, with slopes near −1.

3.3. Agricultural areas

We apply the same progressive segmentation technique to four of the agricultural land cover products to produce rank-size distributions of spatially contiguous segments with 25% and 50% area under cultivation. The resulting distributions, shown in Fig. 10, also produce statistically significant fits to the power law

distribution with slopes near −1 for the two higher resolution products (IIASA-IFPRI and MODIS), but less so for the lower resolution products (Hyde and Earthstat) which show obvious quantization effects in their lower tails. The fits for the higher resolution products yield similar slopes near −1 for both thresholds, suggesting that the structure of the distribution is relatively robust to the choice of threshold.

3.4. Forest cover

We apply the progressive segmentation technique to the MODIS Vegetation Continuous Field tree cover product for a range of tree cover fraction thresholds between 20% and 80% to produce suites of rank-size distributions for each continent. For each distribution, shown in Fig. 11, we obtain best-fit slopes and optimal cutoffs for the power law using the method described above. For each region, except Oceania, we obtain slopes near −1 over a wide range of thresholds. In Europe, North America and Asia the slopes diminish at the highest thresholds indicating that the fragmentation of the largest segments in the upper tail has a more pronounced effect than the attenuation of the smaller segments in the lower tail. The distributions for Africa and South America maintain slopes near −1 over the full range of thresholds used. This is a result of the fact that the forested areas of these continents contain larger, more contiguous areas of dense forest than the other continents and are therefore less subject to fragmentation at higher thresholds. When the distributions of all regions (except Oceania) are combined the resulting global distributions are also well fit by power laws with exponents near −1 for a wide range of thresholds. It is not surprising that forested areas in Oceania do not follow a similar distribution since their areas are constrained by the size distribution of islands produced by a variety of different tectonic and volcanic processes – apparently not well described by the same power law.

4. Discussion

4.1. Land cover as spatial networks

An unexpected, and rather remarkable, result of the comparison of rank-size distributions across land cover category is the tendency for all three to form similar distributions spanning a wide range of scales and simultaneously interspersed within one another on the landscape. The robust similarity of the rank-size distributions of seven independently produced global land cover products seems unlikely to be a coincidence. This suggests (to us)
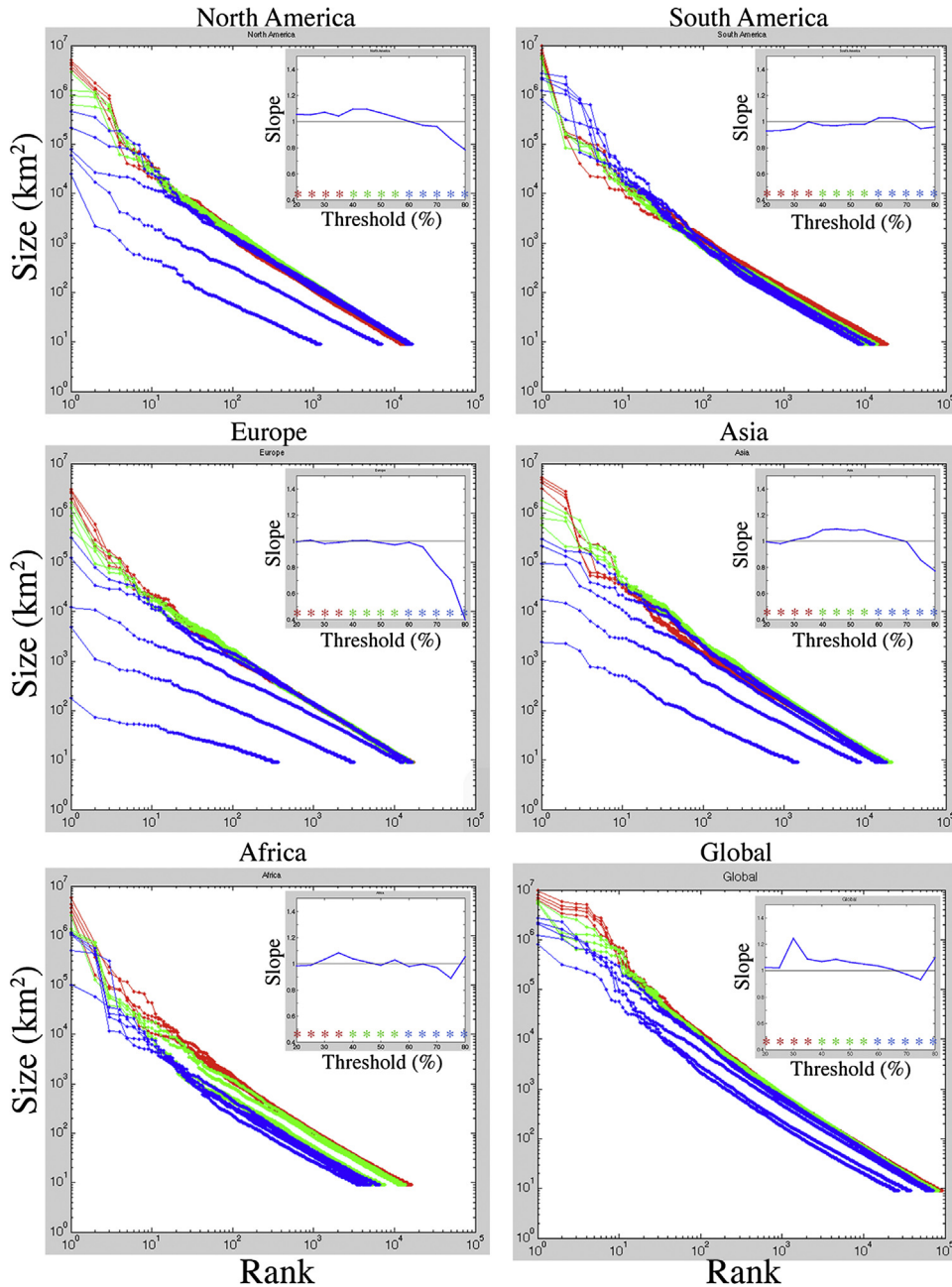
Fig. 11. Rank-size distributions of VCF forest patches in 2000. Increasing tree fraction thresholds reduces the size and connectivity of the spatial networks but has little effect on the slope (power law exponent) – except at the highest fraction thresholds in Europe, Asia and North America. The slope remains near −1 even at the highest thresholds for South America and Africa because both have very large contiguous areas of intact forest with high percent tree cover. At lower thresholds forest patch size distributions have similar slopes on all continents.

that there may be a common explanation that is fundamentally related to spatial structure rather than different process-specific explanations for each land cover type. For this reason, we seek a general explanation for the observed similarity of scaling that does not rely on different processes specific to settlements, agriculture or forests.

One common characteristic that all three types of land cover share is their resemblance to spatial networks. It was observed by Small et al. (2011) that the larger contiguous patches of lighted development do not correspond to individual cities but enormous interconnected spatial networks composed of cities of varying sizes. Similarly, the larger agricultural areas are not composed of

individual farms but contiguous spatial networks of farms of varying size. Because of the similarity of land cover mosaics to spatial networks, and the fact that networks are often characterized by heavy-tailed distributions, we consider the structure of the land cover products as spatial networks. We propose a general representation of land cover as spatial networks occupying a plane, or more generally a surface. In contrast to the convention of representing networks as discrete sets of nodes and links with few, if any, spatial constraints, we consider binary spatial networks that arise from discretization of continuous fields – subject to firm spatial constraints.

A wide variety of other types of networks are characterized by scale-free (e.g. power law) properties (see (Newman, 2010)). In comparison to scale-free networks that are described with power law *degree distributions* (Barabási and Albert, 1999), the bounded spatial networks we observe have power law *component size distributions* in which each spatially contiguous segment is a component of the network. Whereas scale-free networks can emerge from random networks through the combined processes of growth and preferential attachment (Barabási and Albert, 1999), it can be shown that scale-free spatial networks can also emerge from segmented continuous fields through the same combined processes of growth and preferential attachment (Small and Sousa, 2015). The process of progressive segmentation of the continuous field naturally results in growth of a network while the increase in perimeter of the growing components naturally results in preferential attachment to the largest components with the longest perimeters. In other words, the same conditions for emergence of scaling in random networks (growth & preferential attachment) can also explain emergence of scaling on bounded spatial networks. Growth is driven by the process(es) giving rise to the network while preferential attachment is a consequence of growth within the confined domain of the plane or surface. In the case of land cover networks, random growth is sufficient to produce a power law distribution, and any growth process that does not favor a particular size range of component will naturally result in more rapid growth of larger components.

The formation and evolution of a bounded spatial network involves three fundamental processes; nucleation, growth and connection. Each time a new pixel node is added to the domain, one of these three processes occurs. If the new pixel node is isolated and not adjacent to any other occupied pixel nodes we refer to the process as *nucleation*. If the new pixel node appears adjacent to a pixel node that is already part of an existing component we refer to the process as *growth* because the existing node has grown by one pixel. If the new pixel node appears adjacent to two or more pixel nodes that are parts of different components we refer to the process as *connection* because those components become connected into a new larger component. We hypothesize that the parallel evolution of the probabilities of nucleation, growth and connection within the domain controls the emergence of scaling in bounded spatial networks. It can be shown with model simulations that this combination of nucleation, growth and connection can even explain the emergence of scaling on random spatial networks – without the need to introduce land cover-specific processes (Small and Sousa, 2015). Moreover, the processes of nucleation, growth and connection in growing spatial networks have direct complements in the processes of attenuation (nucleation), shrinkage (growth) and fragmentation (connection) in shrinking spatial networks. In binary cases where a domain or landscape is composed of a foreground (e.g. settlements) and a background (e.g. forest), the growth of a foreground spatial network can produce a complementary spatial network on the background whereby the rank-size distributions of both foreground and background networks can be described with power laws at different stages of their evolution. Generalized characteristics for co-evolution of complementary spatial networks are given by Small and Sousa (2015).

The unity power law rank-size distribution is not a general property of all spatial networks derived from progressive segmentation of continuous fields. We have found one prominent exception so far. Applying progressive segmentation to digital elevation models (DEMs) of Earth's continents does not generally yield rank-size distributions with power law upper tails or slopes near −1. Progressive segmentation of 500 m resolution continental DEMs from the USGS yields rank-size distributions with discontinuous upper tails and greater curvature than is observed for the land cover products. Furthermore, it is necessary to use different ranges of elevation thresholds for each continent to obtain heavy-tailed rank size distributions bearing any resemblance to those observed for the land cover products. The slopes ($\mu \pm \sigma$) we obtain from progressive segmentation of DEMs for Asia, Australia, Europe + Africa, North America and South America are -1.49 ± 0.2, -1.43 ± 0.09, -1.48 ± 0.09, -1.39 ± 0.2 and -1.39 ± 0.12 respectively. While the similarity of the exponents is compelling, they are more variable and do not produce the uniform area scaling we observe for land cover networks.

The statistical consistency in the size distributions of settlements, agricultural areas and forests is not characteristic of other distributions commonly described with power laws. In a comparative analysis of 24 different quantities frequently described with power laws, (Clauset et al., 2009), used the same statistical approach to estimate exponents of each. The comparison included quantities as diverse as word usage frequency, book sales, earthquake intensity, city size (population), surname frequency, net worth and war intensity. The median slope of the best fit power laws to these 24 distributions was −0.76 with an interquartile range of 0.37. In comparison, the median slope and interquartile range for the settlement, agricultural area and forest distributions estimated in this study are -0.99 ± 0.05, -1.11 ± 0.21 and -1.05 ± 0.07 respectively. If the two lower resolution (10×10 km) agriculture products are excluded, the median slope and interquartile range for the kilometer-resolution agricultural areas is -1.06 ± 0.07.

The rank-size distributions of kilometer resolution products described here span four to five orders of magnitude in both number and size; a wider range than many distributions commonly described by power laws (see Clauset et al., 2009 for examples). For practical reasons, we limit this analysis to the kilometer-resolution global land cover products. Extending the analysis to the decameter-resolution products is beyond the scope of the present study, but should be done when all of these products have been made publicly available. In the meantime, we have extended the scaling analysis of agricultural land cover to decameter and meter scales using a collection of Landsat, Sentinel and Ikonos imagery from a variety of agricultural areas worldwide (Sousa and Small, 2016). The results of this analysis show that networks of vegetation fraction at decameter and meter scales exhibit substantial variability in spatial distribution, connectivity structure, and robustness to changing threshold. However, the agricultural areas analyzed still possess similar scaling properties to the global distribution of agriculture – despite severely undersampling the complete global agricultural network and representing a wide range of field sizes, competing land uses, climate zones, and land management practices.

We use the power law fits to the rank-size distributions as a convenient tool to quantify the slopes and their degree of similarity – yet we remain agnostic on the question of whether the component distributions of the spatial networks are generated by a strict power law process. As pointed out by Clauset et al. (2009) and others, it can be very difficult to conclusively rule out other heavy-tailed distributions such as lognormal. Plausible arguments have been made for combined power law-lognormal (e.g. (Halloy, 1999)) and double Pareto-lognormal ((Reed, 2002)) distributions for observations generated by complex systems. The goodness-of-fit estimates we obtain are >0.1 in the majority of cases, suggesting that a power law is a plausible, but other heavy-tailed distributions may fit as well or better. We adopt the power law description here for consistency with the large number of studies inferring power laws for other types of networks. While the exponents of the best-fit power laws are close to −1 for the global land cover products, there is some variability in the estimates. Rather than focus on the exact value of the exponent, we

emphasize the physical significance of the slope of −1 as transitional between distributions dominated by larger numbers of small components and those dominated by smaller numbers of large components. In dynamic systems where spatial networks of land cover are continually evolving, the form and slope of the rank-size distribution can provide an indication of the relative importance of nucleation, growth and connection. As such, rank-size distributions of component sizes can provide important constraints for testing hypotheses and models of land use change (e.g. Verburg et al., 1999, 2015) when land cover is depicted as coevolving spatial networks on the landscape.

4.2. Principal challenges of global land cover mapping

For several reasons explained in the Background section, the production of consistent, accurate land cover products at global scales is extremely challenging. In addition to the fundamental challenges related to spatial, spectral and temporal resolution, there are numerous practical challenges not discussed. For optical sensors, the presence and detection of clouds is a difficult challenge in itself. There are many others. While some of these challenges may be resolved in the near future by constellations of sensors with higher spatial, spectral and temporal resolution, there remain conceptual challenges that will be more difficult to resolve. Some of these were alluded to in the discussion of classification algorithms. Identification of consistent physical properties describing the enormous diversity of building materials used in human settlements poses a fundamental challenge for mapping a land cover type that is characterized more by its function than its form (Small, 2009). The identification of impervious surfaces from optical imagery presents a similar challenge. Despite numerous studies attempting to map impervious surfaces with optical imagery, a general solution has yet to be found. In part, this is because permeability is a hydraulic property, not an optical property, so non-uniqueness of reflectance spectra confounds attempts to distinguish permeable soils and substrates from impermeable materials derived from these constituents. Similar challenges exist for mapping agriculture and forests. Multiple scales of geographic variability within each of these classes of land cover can be more consistent than the difference between them across biomes. The fact that agriculture and indigenous vegetation are often interspersed at scales finer than hectometer, or even decameter, resolution, and the fact that leaf level reflectance is frequently not diagnostic between wild and cultivated grasses are indicative of the kind of discrimination problem confronting any attempt to distinguish them consistently – even at regional scales.

4.3. Consistency among land cover products

Despite qualitative differences among products for all three land cover categories, there is a surprising level of agreement among them – particularly when considering that different products attempting to map the same land cover often use different sensors and different approaches and algorithms to map nominally the same quantity. While products in all three categories may appear to differ significantly with respect to the presence, extent and abundance of land cover, the similarity of their rank-size distributions is remarkable. Even in cases where the total mapped area differs considerably, all of their rank-size distributions can be described reasonably well by the same functional form (a power law) with the same fundamental characteristic (the exponent). The similarity of slopes for distributions of independent estimates of the same land cover suggests that each of these land cover products is representing a common fundamental characteristic in the uniformity of area across a wide range of patch sizes –

even if the extents of the patches do not match exactly. Because this descriptor is related to both the overall spatial structure and the process by which each distribution evolves, it suggests (to us) that all the products may be resolving a fundamental characteristic of each of the land cover types.

4.4. Product ensembles

The consistency of the different products in each category of land cover suggests that the data and approaches used to map them are complementary. A comparative accuracy analysis of these products is beyond the scope of this paper, but we expect that the analyses, when eventually conducted on all products, will reveal strengths and limitations of each product. On route to the goal of deriving a single definitive product for each class, a productive approach may follow a strategy similar to that increasingly used to evaluate performance of climate models. Given the different approaches and inputs used in each product, and incomplete knowledge of the process being modeled, it may be informative to evaluate similarities and differences in ensembles of independently derived land cover products. The simple approach used in Figs. 6–8 suggests a strategy for comparing spatial products that map nominally the same land cover type in the same place. Superimposing multiple products would allow for evaluation of both the similarity as the common geographic intersection and the difference as the relative complement between each pair of products. With this approach, the disagreement between products could be at least as informative as the agreement among products. In this sense, the differences between independent products could hold the key to accurate, repeatable mapping and monitoring of anthropogenic land cover using remote sensing.

Land Cover Gallery

A gallery of multi-scale land cover comparisons, similar to those in Fig. 2, is available online at: http://www.LDEO.columbia.edu/~small/LandCoverGallery/.

Further reading

As a complement to the research cited throughout this paper, we provide a selection of additional sources of both knowledge and inspiration on the subject of human modification of Earth's landscapes. All strongly recommended.

Sheffield, C., Man on Earth, 1983, (ISBN-10: 0026101009), Earthwatch, 1981, (ISBN-10: 0026100908) Two of the earliest large format galleries of Landsat imagery.

Johnston, A., Earth from Space, 2004, (ISBN-10: 1552978206) A more recent compendium of visually compelling images from a variety of sensors.

Kelley, K., Cousteau, J.Y., The Home Planet, 1988, (ISBN-10: 0201151979) One of the earliest large format galleries of astronaut photography of Earth.

Gerster, G., Grand Design; Earth from Above, 1976, (ISBN-10: 0846701405) One of the earliest galleries of low altitude color aerial photography of modified landscapes.

Arthus-Bertrand, Y., Earth from Above, 2005, (ISBN-10: 0810996936) A more recent compendium of low altitude color aerial photography of modified landscapes.

Bridges, M. Planet Peru, 1991, (ISBN-10: 0893814695), Markings, 1986, (ISBN-10: 0893812285), Visually compelling collections of black and white aerial photography of terraformed landscapes.

Bacon, E., Design of Cities, 1978, (ISBN-10: 050027133X), A deterministic perspective on urban form and evolution.

Oliver, P., Dwellings, The Vernacular House Worldwide, 1987, (ISBN-10: 0714847933), Comprehensive description of the diversity of human dwellings.

Menzel, P. and Mann C., Material World; A Global Family Portrait, 1994, (ISBN-10: 08715643000), Multicultural perspective on human homes and possessions at the end of the 20th century.

Acknowledgements

References

Adams, J.B., Gillespie, A.R., 2006. Remote Sensing of Landscapes with Spectral Images. Cambridge University Press, Cambridge, UK.
Adams, J.B., Smith, M.O., Johnson, P.E., 1986. Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site. J. Geophys. Res. 91, 8098–8122.
Adams, J.B., Sabol, D.E., Kapos, V., Filho, R.A., Roberts, D.A., Smith, M.O., Gillespie, A.R., 1995. Classification of multispectral images based on fractions of endmembers: application to land-cover change in the Brazilian Amazon. Remote Sens. Environ. 52, 137–154.
Alcantara, C., Kuemmerle, T., Prishchepov, A., Radeloff, V.C., 2012. Mapping abandoned agriculture with multi-temporal MODIS satellite data. Remote Sens. Environ. 124, 334–347.
Angel, S., Sheppard, S., Civco, D., 2005. The Dynamics of Global Urban Expansion. The World Bank, Washington, D.C.
Arino, O., Gross, D., Ranera, F., Leroy, M., Bicheron, P., Brockman, C., Defourny, P., Vancutsem, C., Achard, F., Durieux, L., Bourg, L., Latham, J., Gregorio, A., Witt, R., Herold, M., Sambale, J., Plummer, S., Weber, J.L., 2007. GlobCover: ESA service for global land cover from MERIS. Proceedings of the Geoscience and Remote Sensing Symposium (IEEE International, IGARSS 2007), Barcelona, Spain.
Asner, G.P., Knapp, D.E., Balaji, A., Paez-Acosta, G., 2009. Automated mapping of tropical deforestation and forest degradation: CLASlite. J. Appl. Remote Sens. 3.
Auerbach, F., 1913. Das gesetz der bevolkerungskonzentration. Petermanns Geographische Mitteilungen 59, 74–76.
Balk, D., 2009. More than a name: why is global urban population mapping a GRUMPy proposition? In: Herold, P.G.A.M. (Ed.), Global Mapping of Human Settlement: Experiences, Data Sets, and Prospects. Taylor and Francis.
Barabási, A.-L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.
Barnes, W.L., Pagano, T.S., Salomonson, V.V., 1998. Prelaunch characteristics of the moderate resolution imaging spectroradiometer (MODIS) on EOS-AMI. IEEE Trans. Geosci. Remote Sens. 36, 1088–1100.
Bartholome, E., Belward, S., 2005. GLC2000: a new approach to global land cover mapping from earth observation data. Int. J. Remote Sens. 26, 1959–1977.
Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., Bontemps, S., Leroy, M., Achard, F., Herold, M., Ranera, F., Arino, O., 2008. GlobCover 2005—Products description and validation report. MEDIAS-France. http://publications.jrc.ec.europa.eu/repository/handle/JRC49240.
Biradar, C.M., Thenkabail, P.S., Noojipady, P., Li, Y., Dheeravath, V., Turral, H., Velpuri, M., Gumma, M.K., Gangalakunta, O.R.P., Cai, X.L., Xiao, X., Schull, M.A., Alankara, R.D., Gunasinghe, S., Mohideen, S., 2009. A global map of rainfed cropland areas (GMRCA) at the end of last millennium using remote sensing. Int. J. Appl. Earth Obs. Geoinf. 11, 114–129.
Carfagna, E., Pratesi, M., Carfagna, A., 2013. Methodological developments for improving the reliability and cost-effectiveness of agricultural statistics in developing countries. Proceedings of the 59th World Statistics Congress of the International Statistical Institute (pp. 1930-1935), The Hague, Netherlands: International Statistical Institute.
Carletto, C., Jolliffe, D., Banerjee, R., 2013. The emperor has no data! agricultural statistics in sub-saharan africa. World Bank FAO (2014). FAOSTAT. FAO, Rome, Italy.
CIESIN, 2014. Gridded Population of the World, Version 4 (GPWv4), Preliminary Release 2 (2010). Center for International Earth Science Information Network Earth Institute Columbia University, Palisades NY, USA.
Clauset, A., Shalizi, C.R., Newman, M.E.J., 2009. Power-law distributions in empirical data. SIAM Rev. 51, 661–703.
Cohen, J.E., Small, C., 1998. Hypsographic demography: the distribution of human population by altitude. Proc. Natl. Acad. Sci. U. S. A. 95, 14009–14014.
de Jong, R., Verbesselt, J., Zeileis, A., Schaepman, M., 2013. Shifts in global vegetation activity trends. Remote Sens. 5, 1117–1133.
Deuskar, C., 2015. What does urban mean? In, Sustainable Cities. World Bank Blog. http://blogs.worldbank.org/sustainablecities/what-does-urban-mean.
DiMiceli, C.M., Carroll, M.L., Sohlberg, R.A., Huang, C., Hansen, M.C., Townshend, J.R.G., 2011. Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250 m Spatial Resolution for Data Years Beginning Day 65, 2000–2010 Collection 5 Percent Tree Cover. University of Maryland, College Park, MD, USA.
Dorélien, A., Balk, D., Todd, M., 2013. What is urban? Comparing a satellite view with the demographic and health surveys. Popul. Dev. Rev. 39, 413–439.
Ellis, E.C., Ramankutty, N., 2008. Putting people in the map: anthropogenic biomes of the world. Front. Ecol. Environ. 6, 439–447.
Elvidge, C.D., Baugh, K.E., Kihn, E.A., Kroehl, H.W., Davis, E.R., 1997. Mapping city lights with nighttime data from the DMSP Operational Linescan System. Photogramm. Eng. Remote Sens. 63, 727–734.
Elvidge, C.D., Baugh, K.E., Dietz, J.B., Bland, T., Sutton, P.C., Kroehl, H.W., 1999. Radiance calibration of DMSP-OLS low-Light imaging data of human settlements. Remote Sens. Environ. 68, 77–88.
Elvidge, C.D., Imhoff, M.L., Baugh, K.E., Hobson, V.R., Nelson, I., Safran, J., Dietz, J.B., Tuttle, B.T., 2001. Night-time lights of the world: 1994–1995. ISPRS J. Photogramm. 56, 81–99.
Elvidge, C.D., Ziskin, D., Baugh, K.E., Tuttle, B.T., Ghosh, T., 2009. A fifteen year record of global natural gas flaring derived from satellite data. Energies 2, 595–622.
Elvidge, C.D., Baugh, K.E., Zhizhin, M., Hsu, F.C., 2013. Why VIIRS data are superior to DMSP for mapping nighttime lights. Proc. Asia-Pac. Adv. Netw. 35, 62–69.
Elvidge, C.D., Zhizhin, M., Baugh, K., Hsu, F.-C., 2015. Automatic boat identification system for VIIRS low light imaging data. Remote Sens. 7, 3020–3036.
Esch, T., Schenk, A., Ullmann, T., Thiel, M., Roth, A., Dech, S., 2011. Characterization of land cover types in TerraSAR-X images by combined analysis of speckle statistics and intensity information. IEEE Trans. Geosci. Remote Sens. 49 (6), 1911–1925.
Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenböck, H., Müller, A., Dech, S., 2013. Urban footprint processor—fully automated processing chain generating settlement masks from global data of the tanDEM-X mission. IEEE Trans. Geosci. Remote Sens. 10, 1617–1621.
Fensholt, R., Langanke, T., Gaughan, A.E., Rasmussen, K., Reenberg, A., Paruelo, J., Tucker, C., Eastman, R., Le, Q.B., Wessels, K., Scholes, R.J., Mbow, C., Seaquist, J., Hellden, U., Prince, S.D., Bondeau, A., Olsson, L., Schweitzer, C., Epstein, H., 2012. Greenness in semi-arid areas across the globe 1981–2007—An Earth Observing Satellite based analysis of trends and drivers. Remote Sens. Environ. 121, 144–158.
Foley, J.A., Ramankutty, N., Brauman, K.A., 2011. Solutions for a cultivated planet. Nature 478, 337–342.
Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., Schaaf, C., 2002. Global land cover mapping from MODIS: algorithms and early results. Remote Sens. Environ. 83, 287–302.
Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., 2009. MODIS collection 5 global land cover: algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182.
Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168–182.
Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., Duerauer, M., Albrecht, F., Schill, C., Perger, C., Havlik, P., Mosnier, A., Thornton, P., Wood-Sichra, U., Herrero, M., Becker-Reshef, I., Justice, C., Hansen, M., Gong, P., Abdel Aziz, S., Cipriani, A., Cumani, R., Cecchi, G., Conchedda, G., Ferreira, S., Gomez, A., Haffani, M., Kayitakire, F., Malanding, J., Mueller, R., Newby, T., Nonguierma, A., Olusegun, A., Ortner, S., Ramrajak, D., Rocha, J., Schepaschenko, D., Schepaschenko, M., Terekhov, A., Tiangwa, A., Vancutsem, C., Vintrou, E., Wenbin, W., Van der Velde, M., Dunwoody, A., Kraxner, F., Obersteiner, M., 2015. Mapping global cropland and field size. Glob. Change Biol. 21, 1980–1992.
Gabaix, X., Ioannides, Y.M., Henderson, J.V., Jacques-Fránçois, T., 2004. Chapter 53 The evolution of city size distributions. Handbook of Regional and Urban Economics. Elsevier, pp. 2341–2378.
Giri, C., Pengra, B., Zhu, Z., Singh, A., Tieszen, L.L., 2007. Monitoring mangrove forest dynamics of the Sundarbans in Bangladesh and India using multi-temporal satellite data from 1973 to 2000. Estuarine, Coastal Shelf Sci. 73, 91–100.
Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., Li, C., Li, X., Fu, W., Liu, C., Xu, Y., Wang, X., Cheng, Q., Hu, L., Yao, W., Zhang, H., Zhu, P., Zhao, Z., Zhang, H., Zheng, Y., Ji, L., Zhang, Y., Chen, H., Yan, A., Guo, J., Yu, L., Wang, L., Liu, X., Shi, T., Zhu, M., Chen, Y., Yang, G., Tang, P., Xu, B., Giri, C., Clinton, N., Zhu, Z., Chen, J., Chen, J., 2013. Finer resolution observation and monitoring of

global land cover: first mapping results with Landsat TM and ETM+ data. Int. J. Remote Sens. 34, 2607–2654.

Halloy, S.R.P., 1999. A Theoretical Framework for Abundance Distributions in Complex Systems. Santa Fe Institute, Santa Fe, NM.

Creation of a high spatiotemporal resolution global database of continuous mangrove forest cover for the 21st Century (CGMFC-21): A big-data fusion approach. arXiv http://arxiv.org/abs/1412.0722http://arxiv.org/abs/1412.0722

Hansen, M.C., Defries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. Int. J. Remote Sens. 21, 37–41.

Hansen, M.C., DeFries, R.S., Townshend, J.R.G., Sohlberg, R., Dimiceli, C., Carroll, M., 2002. Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data. Remote Sens. Environ. 83, 303–319.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853.

in, Houghton, J.T., Filho, L.G.M., Callander, B.A., Harris, N., Kattenberg, A., Maskell, K. (Eds.), 1996. Climate Change 1995. Cambridge University Press, Cambridge UK.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213.

Iizumi, T., Yokozawa, M., 2014. Historical changes in global yields: major cereal and legume crops from 1982 to 2006. Glob. Ecol. Biogeogr. 23, 346–357.

Ioannides, Y.M., Overman, H.G., 2003. Zipf's law for cities: an empirical examination. Reg. Sci. Urban Econ. 33, 127–137.

Jensen, J.R., 2007. Remote Sensing of the Environment; An Earth Resource Perspective. Pearson Prentice Hall, Upper Saddle River, NJ.

Kempeneers, P., Sedano, F., Pekkarinen, A., Seebach, L., Strobl, P., San-Miguel-Ayanz, J., 2012. Pan-European forest maps derived from optical satellite imagery. IEEE Earthzine 5, 390004.

Kim, D.-H., Sexton, J.O., Noojipady, P., Huang, C., Anand, A., Channan, S., Feng, M., Townshend, J.R., 2014. Global, Landsat-based forest-cover change from 1990 to 2000. Remote Sens. of Environ. 155, 178–193.

Klein Goldewijk, K., Van Drecht, G., Bouwman, A.F., 2006. Mapping contemporary global cropland and grassland distributions on a 5 × 5 minute resolution. J. Land Use Sci. 2, 167–190.

Klein Goldewijk, K., Verburg, P.H., 2013. Uncertainties in global-scale reconstructions of historical land use: an illustration using the HYDE data set. Landscape Ecol. 28, 861–877.

Klotz, M., Kemper, T., Geiß, C., Esch, T., Taubenböck, H., 2016. How good is the map? A multi-scale cross-comparison framework for global settlement layers: evidence from Central Europe. Remote Sens. Environ. 178, 191–212.

Krieger, G., Moreira, A., Fiedler, H., Hajnsek, I., Werner, M., Younis, M., Zink, M., 2007. TanDEM-X: A satellite formation for high resolution SAR interferometry. IEEE Trans. Geosci. Remote Sens. 45, 3317–3341.

Leff, B., Ramankutty, N., Foley, J.A., 2004. Geographic distribution of major crops across the world. Glob. Biogeochem. Cycles 18.

Li, X., Li, D., 2014. Can night-time light images play a role in evaluating the Syrian Crisis? Int. J. Remote Sens. 35, 6648–6661.

Linard, C., Gilbert, M., Tatem, A.J., 2011. Assessing the use of global land cover data for guiding large area population distribution modelling. Geojournal 76, 525–538.

Lotka, A., 1941. The law of urban concentration. Science 94, 164.

Mertes, C.M., Schneider, A., Sulla-Menashe, D., Tatem, A.J., Tan, B., 2015. Detecting change in urban areas at continental scales with MODIS dat. Remote Sens. Environ. 158, 331–347.

Miller, S.D., Straka, W., Mills, S.P., Elvidge, C.D., Lee, T.F., Solbrig, J., Walther, A., Heidinger, A.K., Weiss, S.C., 2013. Illuminating the capabilities of the suomi national polar-orbiting partnership (NPP) visible infrared imaging radiometer suite (VIIRS) day/night band. Remote Sens. 5, 6717–6766.

Monfreda, C., 2008. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Glob. Biogeochem. Cycles 22.

New, M.G., Hulme, M., Jones, P.D., 1999. Representing twentieth-century space-time climate variability pt. 1. Development of a 1961- mean monthly terrestrial climatology. J. Clim. 12, 829–856.

Newman, M.E., 2010. Networks: an Introduction. University Press Scholarship Online, Oxford, England: Oxford.

Nitsch, V., 2005. Zipf zipped. J. Urban Econ. 57, 86–100.

Overman, H.G., Ioannides, Y.M., 2001. Cross-sectional evolution of the U.S: city size distribution. J. Urban Econ. 49, 543–566.

Pesaresi, M., Benediktsson, J., 2001. A new approach for the morphological segmentation of high-resolution satellite imagery. IEE Trans. Geosci. Remote Sens. 39, 309–320.

Pesaresi, M., Gerhardinger, A., Kayitakire, F., 2008. A robust built-up area presence index by anisotropic rotation-invariant textural measure. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 1, 180–192.

Pesaresi, M., Ehrlich, D., Ferri, S., Florczyk, A., Freire, S., Halkia, M., Julea, A., Kemper, T., Soille, P., Syrris, V., 2015. A new map of the global human settlements from 40-years record of Landsat satellite data. Remote Sens. Environ. (Under Review).

Pittman, K., Hansen, M.C., Becker-Reshef, I., Potapov, P.V., Justice, C.O., 2010. Estimating global cropland extent with multi-year MODIS data. Remote Sens. 2, 1844–1863.

Potapov, P., Turubanova, S., Hansen, M.C., 2011. Regional-scale boreal forest cover and change mapping using Landsat data composites for European Russia. Remote Sens. Environ. 115, 548–561.

Potere, D., Schneider, A., 2009. Comparison of global urban maps. In: Herold, P.G.M. (Ed.), Global Mapping of Human Settlement. Experiences, Datasets, and Prospects. CRC Press, Boca Raton, FL, pp. 269–308.

Pumain, D., 2004. Scaling Laws and Urban Systems. Santa Fe Institute, Santa Fe NM, pp. 26.

Ramankutty, N., Evan, A.T., Monfreda, C., Foley, J.A., 2000. Farming the planet: 1. geographic distribution of global agricultural lands in the year. Glob. Biogeochem. Cycles 22.

Rast, M., Bezy, J.L., Bruzzi, S., 1999. The ESA medium resolution imaging spectrometer MERIS—a review of the instrument and its mission. Int. J. Remote Sens. 20, 1681–1702.

Reed, W.J., 2002. On the rank-size distribution for human settlements. J. Reg. Sci. 42, 1–17.

Rose, A.N., Bright, E., 2014. The LandScan Global Population Distribution Project: Current State of the Art and Prospective Innovation. Population Association of America, Boston MA (Population Association of America).

Rosen, K.T., Resnick, M., 1980. The size distribution of cities: an examination of the Pareto law and primacy. J. Urban Econ. 8, 165–186.

Salmon, J.M., Friedl, M.A., Frolking, S., Wisser, D., Douglas, E.M., 2015. Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. Int. J. Appl. Earth Obs. Geoinf. Int. J. Appl. Earth 38, 321–334.

Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N., Zhang, X.Y., Jin, Y., Muller, J.P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R.P., Hu, B., Liang, S., Privette, J.L., 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sens. Environ. 83, 135–149.

Schepaschenko, D., McCallum, I., Shvidenko, A., Fritz, S., Kraxner, F., Obersteiner, M., 2011. A new hybrid land cover dataset for Russia: a methodology for integrating statistics, remote sensing and in situ information. J. Land Use Sci. 6, 245–259.

Schepaschenko, D., See, L., Lesiv, M., McCallum, I., Fritz, S., Salk, C., Moltchanova, E., Perger, C., Shchepashchenko, M., Shvidenko, A., Kovalevskyi, S., Gilitukha, D., Albrecht, F., Kraxner, F., Bun, A., Maksyutov, S., Sokolov, A., Dürauer, M., Obersteiner, M., Karminov, V., Ontikov, P., 2015. Development of a global hybrid forest mask through the synergy of remote sensing, crowdsourcing and FAO statistics. Remote Sens. Environ. 162, 208–220.

Schneider, A., Friedl, M.A., Potere, D., 2010. Monitoring global urban areas using MODIS 500 m data: new methods and datasets based on urban ecoregions. Remote Sens. Environ. 114, 1733–1746.

Schneider, A., Mertes, C.M., Tatem, A.J., Tan, B., Sulla-Menashe, D., Graves, S.J., Patel, N.N., Horton, J.A., Gaughan, A.E., Rollo, J.T., Schelly, I.H., Stevens, F.R., Dastur, A., 2015. A new urban landscape in East–Southeast Asia, 2000–2010. Environ. Res. Lett. 10.

Seebach, L., McCallum, I., Fritz, S., Kindermann, G., Leduc, S., Böttcher, H., Fuss, S., 2012. Choice of forest map has implications for policy analysis: a case study on the EU biofuel target. Environ. Sci. Policy 22, 13–24.

Seto, K.C., Fragkias, M., Guneralp, B., Reilly, M.K., 2011. A meta-analysis of global urban land expansion. PLoS One 6.

Sexton, J.O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K.M., Channan, S., Dimiceli, C., Townshend, J.R., 2013. Global, 30 m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. Int. J. Digit. Earth 6, 427–448.

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., Lucas, R., 2014. New global forest/non-forest maps from ALOS PALSAR data. Remote Sens. Environ. 155, 13–31.

Siebert, S., Döll, P., Hoogeveen, J., Faures, J.-M., Frenken, K., Feick, S., 2005. Development and validation of the global map of irrigation areas. Hydrol. Earth Syst. Sci. (Discussions, 2, 1299–1327).

Skole, D.L., Tucker, C.J., 1993. Tropical deforestation and habitat fragmentation in the Amazon—satellite data from 1978 to 1988. Science 260, 1905–1910.

Skole, D.L., Cochrane, M.A., Matricardi, E., Chomentowski, W.H., Pedlowski, M., Kimble, D. (Eds.), 2004. Pattern to Process in the Amazon Region: Measuring Forest Conversion, Regeneration and Degradation. Kluwer Academic Publishers.

Small, C., Cohen, J.E., 2004. Continental physiography, climate, and the global distribution of human population. Curr. Anthropol. 45, 269–277.

Small, C., Elvidge, C.D., 2013. Night on Earth: mapping decadal changes of anthropogenic night light in Asia. Int. J. Appl. Earth Obs. Geoinf. 22, 40–52.

Small, C., Sousa, D., 2015. Spatial Scaling of Land Cover Networks. ArXiv, pp. 1–23. http://arxiv.org/abs/1512.01517.

Small, C., Montgomery, M., Balk, D., Elvidge, C.D., Bright, E.A., Yetman, G., 2011. Global Scaling and Connectivity of Human Population and Development. ArXiv.

Small, C., 2009. The color of cities: an overview of urban spectral diversity. In: Herold, M., Gamba, P. (Eds.), Global Mapping of Human Settlements. Taylor & Francis, CRC.

Small, C., 2012. Spatiotemporal dimensionality and time-space characterization of multitemporal imagery. Remote Sens. Environ. 124, 793–809.

Smith, M.O., Ustin, S.L., Adams, J.B., Gillespie, A.R., 1990. Vegetation in deserts: I: A regional measure of abundance from multispectral images. Remote Sens. Environ. 31, 1–26.

Soo, K.T., 2005. Zipf's law for cities: a cross-country investigation. Reg. Sci. Urban Econ. 35, 239–263.

Sornette, D., 2003. Critical Phenomena in Natural Science, 2nd ed. Springer-Verlag, Heidelberg.

Sousa, D., Small, C., 2016. Spatial Structure and Scaling of Agricultural Networks. ArXiv. https://arxiv.org/submit/1535732.

Stevenson, R., Sarker, A., May, M.A. (2013). Frugal map making: experiences from Korail. In, BRAC blog. http://blog.brac.net/2013/09/frugal-map-making-experiences-from-korail/.

Stevens, F.R., Gaughan, A.E., Linard, C., Tatem, A.J., 2015. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. PLoS One 10, 1–22 e0107042.

Tateishi, R., Uriyangqai, B., Al-Bilbisi, H., Ghar, M.A., Tsend-Ayush, J., Kobayashi, T., Kasimu, A., Hoan, N.T., Shalaby, A., Alsaaideh, B., Enkhzaya, T., Gegentana, Sato, HP, 2011. Production of global land cover data ? GLCNMO. Int. J. Digit. Earth 4, 22–49.

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., Dech, S., 2012. Monitoring urbanization in mega cities from space. Remote Sens. Environ. 117, 162–176.

Thenkabail, P.S., Biradar, C., Noojipady, P., Dheeravath, V., Li, Y., Velpuri, M., Gumma, M., Gangalakunta, O., Turral, H., Cai, X., Vithanage, J., Schull, M., Dutta, R., 2009. Global irrigated area map (GIAM), derived from remote sensing: for the end of the last millennium. Int. J. Remote Sens. 30, 3679–3733.

Tucker, C.J., Grant, D.M., Dykstra, J.D., 2004. NASA's global orthorectified Landsat data set. Photogramm. Eng. Remote Sens. 70, 313–322.

Verburg, P.H., Koning, G.D., Kok, K., Veldkamp, A., Bouma, J., 1999. A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. Ecol. Model. 116, 45–61.

Verburg, P.H., Dearing, J.A., Dyke, J.G., Leeuw, S.v.d, Seitzinger, S., Steffen, W., Syvitski, J., 2015. Methods and approaches to modelling the Anthropocene. Global Environ. Change (in press).

Verhegghen, A., Mayaux, P., de Wasseige, C., Defourny, P., 2012. Mapping Congo Basin forest types from 300 m and 1km multi-sensor time series for carbon stocks and forest areas estimation. Biogeosciences 9, 5061–5079.

Vijayaraj, V., Bright, E.A., Bhaduri, B.L., 2007. High resolution urban feature extraction for global population mapping using high performance computing. Geosci. Remote Sens. Symp. 278–281 (IEEE International).

Vitousek, P.M., Mooney, H.A., Lubchenco, J., Melillo, J.M., 1997. Human domination of Earth's ecosystems. Science 277, 494–499.

Welch, R., 1980. Monitoring urban-population and energy-utilization patterns from satellite data. Remote Sens. Environ. 9, 1–9.

You, L., Wood, S., Wood-Sichra, U., Wu, W., 2014. Generating global crop distribution maps: from census to grid. Agric. Syst. 127, 53–60.

Yu, L., Wang, J., Clinton, N., Xin, Q., Zhong, L., Chen, Y., Gong, P., 2013. FROM-GC: 30 m global cropland extent derived through multisource data integration. Int. J. Digit. Earth 6, 521–533.

Zhao, Y., Gong, P., Yu, L., Hu, L., Li, X., Li, C., Zhang, H., Zheng, Y., Wang, J., Zhao, Y., Cheng, Q., Liu, C., Liu, S., Wang, X., 2014. Towards a common validation sample set for global land-cover mapping. Int. J. Remote Sens. 35, 4795–4814.

Zipf, G.K., 1942. The unity of nature, least-action and natural social science. Sociometry 5, 48–62.

Zipf, G.K., 1949. Human Behavior and the Principle of Least-Effort. Wesley, Addison.