Mitig Adapt Strateg Glob Change (2017) 22:947–972
DOI 10.1007/s11027-016-9709-9

ORIGINAL ARTICLE

A comparison of five high-resolution spatially-explicit, fossil-fuel, carbon dioxide emission inventories for the United States

Maya G. Hutchins[1,2,3] · **Jeffrey D. Colby**[2] · **Gregg Marland**[4] · **Eric Marland**[5]

Received: 26 October 2015 / Accepted: 23 February 2016 / Published online: 31 March 2016
© Springer Science+Business Media Dordrecht 2016

Abstract The quantification of fossil-fuel-related emissions of carbon dioxide to the atmosphere is necessary in order to accurately represent carbon cycle fluxes and to understand and project the details of the global carbon cycle. In addition, the monitoring, reporting, and verification (MRV) of carbon dioxide emissions is necessary for the success of international agreements to reduce emissions. However, existing fossil-fuel carbon dioxide ($FFCO_2$) emissions inventories vary in terms of the data and methods used to estimate and distribute $FFCO_2$. This paper compares how the approaches used to create spatially explicit $FFCO_2$ emissions inventories affect the spatial distribution of emissions estimates and the magnitude of emissions estimates in specific locales. Five spatially explicit $FFCO_2$ emission inventories were compared: Carbon Dioxide Information and Analysis Center (CDIAC), Emission Database for Global Atmospheric Research (EDGAR), Fossil Fuel Data Assimilation System (FFDAS), Open-source Data Inventory for Anthropogenic CO_2 (ODIAC), and Vulcan. The effects of using specific data and approaches in the creation of spatially explicit $FFCO_2$ emissions inventories, and the effect of resolution on data representation are analyzed using graphical, numerical, and cartographic approaches. We examined the effect of using top-down versus bottom-up

Electronic supplementary material The online version of this article (doi:10.1007/s11027-016-9709-9) contains supplementary material, which is available to authorized users.

✉ Maya G. Hutchins
mghutchi@asu.edu

[1] School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ, USA

[2] Department of Geography and Planning, Appalachian State University, Boone, NC, USA

[3] School of Life Sciences, Arizona State University, 427 East Tyler Mall, Maya Hutchins, LSA 369, Tempe, AZ 85287, USA

[4] Research Institute for Environment, Energy, and Economics, Appalachian State University, Boone, NC, USA

[5] Department of Mathematical Sciences, Appalachian State University, Boone, NC, USA

approaches, nightlights versus population proxies, and the inclusion of large point sources. The results indicate that the approach used to distribute emissions in space creates distinct patterns in the distribution of emissions estimates and hence in the estimates of emissions in specific locations. The different datasets serve different purposes but collectively show the key role of large point sources and urban centers and the strong relationship between scale and uncertainty.

Keywords Carbon dioxide · Emissions inventories · Emissions uncertainty · Geographic information systems · Spatial analysis · United States (US) CO_2 emissions

1 Introduction

Emissions of carbon dioxide (CO_2) from combustion of fossil fuels and some industrial processes, such as the calcination of limestone to produce cement, are resulting in increased levels of CO_2 in the atmosphere and consequently driving changes in the Earth's climate system. By 2011 the global atmospheric concentration of CO_2 had increased by 40 % over its 1750, pre-industrial, value (Hartmann et al. 2013). Estimates of emissions at the global and national level are now widely reported and have become important elements of public policy and mitigation efforts. For the details of geochemical understanding and the monitoring and verification of agreements at all levels, it is becoming increasingly important to estimate emissions at finer spatial and temporal scales. In response, there are now multiple attempts to provide fossil fuel carbon dioxide ($FFCO_2$) emissions inventories at different scales and resolutions.

Fine resolution emissions data are used particularly in inverse modeling, where atmospheric concentrations and transport data are used to derive estimates of emissions sources and sinks, and a priori source estimates can be used to improve final estimates. Ultimately, we are interested in fine resolution, both spatially and temporally, for not only inverse modeling but also for future calibration with high-resolution satellite monitoring of atmospheric concentrations of $FFCO_2$. In particular, this paper focuses on the spatial distribution of emissions estimates because future monitoring and verification of $FFCO_2$, through inverse modeling and high-resolution satellite monitoring, will rely on the accurate spatial inventory of $FFCO_2$ emissions estimates. This paper investigates how the different approaches used in the creation of spatially explicit $FFCO_2$ emissions inventories affect the spatial distribution of emissions estimates at subnational scales and the emissions magnitudes at specific locales. In addition, the effect of scale and resolution on the representation of $FFCO_2$ emissions is examined. This paper uses the two meanings of scale most related to gridded digital data: geographic scale (extent) and measurement scale (resolution) (Cao and Lam 1997; Goodchild 2001, 2011). Specifically, the magnitude and spatial distribution of five spatially explicit $FFCO_2$ data sets are compared for the continental United States (US) at a range of scales using a geographic information system (GIS). The spatially explicit $FFCO_2$ emissions inventories analyzed include the Carbon Dioxide Information and Analysis Center (CDIAC) dataset, the Emission Database for Global Atmospheric Research (EDGAR) dataset, the Fossil Fuel Data Assimilation System (FFDAS) dataset, the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC), and Vulcan (see Table 1). $FFCO_2$ emissions inventories were analyzed for the years 2002 and 2008 due to the limited temporal scale of Vulcan, available only for the year 2002, and ODIAC, available only for the year 2008. This comparison was focused on the US because, whereas the other four top-down data sets are global, the unique, bottom-up Vulcan data set is available only for the US. There are additional high-resolution

Table 1 An overview of FFCO2 emissions inventories analyzed

	CDIAC	EDGAR	FFDAS	ODIAC	Vulcan
Years analyzed	2002, 2008	2002, 2008	2002, 2008	2008	2002
Approach	Top- down	Hybrid	Hybrid	Hybrid	Bottom- up
Version	V2013	V4.2 FT2010	6/28/2014	2/6/2014	V2.2
Energy stats.	U.N.	IEA, U.N.,Commercial	IEA	CDIAC	EPA, NMIM, Aero2k
Scale and resolution	Global, 0.1 °	Global, 1°	Global, 0.1 °	Global, 1 km	USA, 0.1 °
Websites	http://cdiac.ornl.gov/	http://edgar.jrc.ec.europa.eu	http://gurney.faculty.asu.edu/research/ffdas.php	http://odiac.org/dataset	http://vulcan.project.asu.edu/
Large-point sources	N/A	Included	Included	CARMA	NEI and ETS/CEMS
Sectors	Fossil fuel combustion, cement manufacture, gas flaring in oil fields	Energy, industrial processes, solvent and other product use	Energy, manufacturing, industrial, transport, other (residential, agriculture, fishing)	Fossil fuel combustion, cement manufacture, gas flaring in oil fields	Air, cement, commercial, residential, industrial, cement manufacture, utilities, on-road, and non-road
Spatial distribution	Population	Location of LPS, road network, shipping and aviation routes, population density	Nightlights, population, GDP, energy intensity, carbon intensity	Nightlights and location of large point sources	Road networks, location of large point sources, total sq. floor footage of buildings

data sets with less than global scope that could be useful for similar comparisons in other areas (see, for example, Vogel et al. 2013, for Europe).

Detailed methods for estimating the spatial distribution of $FFCO_2$ emissions vary and no comprehensive comparison of the resulting datasets has previously been performed. Such comparisons are necessary to understand the importance of using different data and approaches in the creation of spatially explicit $FFCO_2$ emissions inventories. Additionally, issues of scale and resolution are traditionally important issues in geography, and as GISs have advanced, multi-scale data is starting to play a more important role in studies such as global change (Cao and Lam 1997). Both geographic scale and measurement scale have large implications for the development and use of $FFCO_2$ emissions inventories. This analysis evaluates the difference between more time consuming methods like those used in bottom-up inventories and less detailed top-down approaches. Comparisons of the different approaches will inform future development of gridded distributions of $FFCO_2$ emissions and clarify their associated uncertainty, with the ultimate goal of creating detailed yet globally consistent $FFCO_2$ emissions inventories. A variety of graphical and numerical methods were employed to compare the existing, spatially explicit $FFCO_2$ data sets and to explore how their methods and selection of proxy data are reflected in their final products.

Comparisons between datasets were conducted using metrics found in similar analyses of $FFCO_2$ datasets (e.g., Andres et al. 1996, 2011; Marland et al. 1999; Gregg and Andres 2008; Gurney et al. 2009; Rayner et al. 2010; Oda and Maksyutov 2011; Asefi-Najafabady et al. 2014) including spatial correlation, sum of absolute differences, and difference maps. Cumulative emissions curves, distribution curves, and spatial distribution maps were also analyzed to gain greater insight into the relationships among the multiple emissions inventories. The data sets were compared at various levels of aggregation (0.1°, 0.5°, 1°, 2°, and 3°) to assess how the virtues of each emissions inventory can best be utilized.

1.1 A brief overview

Fossil fuel carbon dioxide ($FFCO_2$) emissions are historically estimated at national and annual scales from national level energy statistics published by either the International Energy Agency (IEA), the United Nations (UN), British Petroleum Corporation (BP), the U.S. Department of Energy/Energy Information Administration (DOE/EIA), or the U.S. Environmental Protection Agency (EPA) (see Table 1). Energy statistics are often divided into sectors and/or fuel types, and emissions per fuel type are calculated from emissions factors, the ratio of CO_2 emitted per unit of fossil fuel burned. There are now five widely used emissions inventories that use one or more of the national level energy statistics data sets published by the above agencies to estimate CO_2 emissions at the national level: IEA, CDIAC, USDOE/EIA, BP, and EDGAR (see Table 1). All of these emissions inventories are global and annual in scale with national level data. However, the "emphases, categories, units, unit conversions and reporting, data processing, and quality assurance" of energy statistics vary across organizations, leading to variability in the estimated magnitudes of emissions (Andres et al. 2012).

1.2 Fine resolution emissions inventories

Despite growing concerns for the impacts of increasing concentrations of CO_2 in the atmosphere, little work has been done to estimate the seasonal or finer temporal flux or the subnational geographic distribution of carbon dioxide emissions (Gregg and Andres 2008).

As a result, modern climate change research is limited in its ability to project the future behavior of the carbon cycle and to provide informative assessments of biogeochemical feedbacks within the carbon cycle; i.e. to project how the natural carbon sinks in the ocean and terrestrial biosphere will perform into the future (Gurney et al. 2009; Oda and Maksyutov 2011; Andres et al. 2011). Fossil fuel consumption statistics have been historically compiled at national extents because energy data were generally available at large scales and questions regarding global climate change during the twentieth century did not require data at smaller scales. Recent climate and carbon science, as well as public policy, necessitates an increase in the spatio-temporal resolution of $FFCO_2$ emissions inventories (Gurney et al. 2009; Andres et al. 2011; Peylin et al 2011).

Finer resolution $FFCO_2$ emissions inventories seek to estimate the location of $FFCO_2$ emissions at the subnational scale (extent) at resolutions of 1° or finer. $FFCO_2$ estimates are used as a priori information in carbon flux models to distinguish between human and natural sources of CO_2, and supporting evidence for anthropogenic contributions to CO_2 emissions relies on "bottom-up" inventories of $FFCO_2$ emissions (Rayner et al. 2010; Andres et al. 2011). Peylin et al. (2011) have shown that "although fossil-fuel emissions are often considered as a 'well-known' term in the terrestrial carbon balance…" temporal and spatial differences between emissions inventories are "critical when comparing models to atmospheric measurements." The Peylin et al. (2011) European study shows that although transport model uncertainties are the largest contributor to current uncertainty, it is necessary to improve "finer spatial and temporal scales…to aid regional modeling studies."

More specifically, spatially distributed $FFCO_2$ emissions are necessary to identify areas of pollution concentration for inputs into air simulation models and result in decreased errors in dispersion emissions transport models (Dai and Rocke 2000; Wang et al. 2012). Furthermore, spatially explicit $FFCO_2$ emissions inventories lend insight into the relationship between fossil fuel use and economic strength (Andres et al. 2011). By improving the spatial and temporal resolution of bottom-up emissions inventories it may be possible to corroborate atmospheric composition with remotely sensed data, making confirmation of international emissions estimates possible (Andres et al. 2011). The magnitude and spatial distribution of $FFCO_2$ emissions at fine resolutions are necessary for the monitoring, reporting, and verification (MRV) of carbon emissions, and will become increasingly important as cap and trade programs or other emissions mitigation programs are developed and the global community continues to move toward a global emissions agreement.

In response to the need for subnational data on $FFCO_2$ emissions there has been increasing interest in creating spatially explicit $FFCO_2$ emissions inventories at resolutions of 1° or finer. Subnational resolution emissions inventories are spatially distributed using three primary methodologies (top-down approaches, bottom-up approaches, and hybrid approaches) and two primary data types (from non-point and point sources). In this analysis we make a distinction between emissions from small point sources and large point sources (LPS). Small point sources of emissions, such as homes, cars, and small businesses, are statistically treated as areal (non-point) sources in emissions inventories. Large point source emissions describe emissions from large facilities related to the generation of electric power, industrial processes, and petroleum refining. However, data are not generally available to characterize emissions from industrial facilities and refineries at the global scale. The term LPS in this paper thus refers strictly to emissions from large electric power generating facilities.

Non-point source $FFCO_2$ emissions estimates are derived from the national level CO_2 emissions inventories referenced previously (IEA, CDIAC, USDOE/EIA, BP, and EDGAR).

However, some emissions inventories, such as CDIAC, do not distinguish between non-point and point source emissions, while other emissions inventories such as EDGAR do. Non-point source data can be disaggregated from the national scale to subnational scales and finer resolutions using top-down approaches. Top-down approaches generally utilize proxies such as population density and/or satellite-observed nightlights data to spatially distribute emissions within a country (Andres et al. 1996, 2011; Gurney et al. 2009; Rayner et al. 2010; Oda and Maksyutov 2011; Wang et al. 2012). The resolution of spatially explicit $FFCO_2$ emissions inventories that utilize a top-down approach rely on the resolution of the proxy data being used. Bottom-up approaches generally involve data collection on fuel consumption or emissions at the building or highway segment scale or lower and sum individual sources of emissions to estimate $FFCO_2$ emissions at the county-, state-, or national scales (Gurney et al. 2009). Hybrid approaches spatially distribute some emissions in space using proxies from the top-down approach, such as nightlights and populations, in conjunction with individual sources used in bottom-up approaches, such as LPS and road networks. For example, data assimilation uses observational data to constrain dynamic models, such as models based on the Kaya identity, (which estimates $FFCO_2$ emissions using data on population, gross domestic product (GDP), and energy consumption (Rayner et al. 2010)). In the Results and discussion section, the term top-down will be used to refer to both traditional top-down approaches and hybrid approaches.

Emissions from large point sources account for approximately 40 % of total national emissions in most countries (Singer et al. 2014). As a result, emissions from LPS data have been recently utilized in the development of spatially explicit $FFCO_2$ emissions inventories and have been incorporated into top-down, bottom-up, and hybrid approaches. Emissions from LPS of $FFCO_2$ are estimated using material balance calculations on fuel consumption and/or continuous emissions monitors (CEMs). Material balance approaches multiply the difference between the amount of material entering a process and the final product by an emissions factor to estimate the CO_2 emissions that result from a specific process. In countries without data on emissions from LPS, regression analysis has been used to estimate emissions (Wheeler and Ummel 2008). LPS emissions data are available from the U.S. EPA, the U.S. DOE/EIA, and the Center for Global Development. The U.S. EPA publishes three data sets on LPS emissions, the Emissions & Generation Resource Integrated Database (eGRID), the National Emissions Inventory (NEI), and the Greenhouse Gas Reporting Program (GHGRP). The GHGRP and NEI include emissions estimates from all LPSs, while eGRID estimates CO_2 emissions from electrical generation only. The NEI reports emissions of carbon monoxide (CO) from point sources in the U.S., which can be converted to estimate CO_2 emissions using emissions factors. The Center for Global Development publishes CARMA (Carbon Monitoring for Action), which reports emissions of CO_2 from electric power plants globally.

2 Data and methods

2.1 Data

Materials used in this analysis include five data sets on gridded $FFCO_2$ emissions inventories for the USA. A comparison is made among the five $FFCO_2$ emissions inventories for similarities and differences in the energy statistics used, the sectors included and, most significantly, the methods used for spatial disaggregation or allocation of emissions

(Table 1). Reliance on data from point or line sources varies as does the proxy for distributing emissions that are not attributable to point or line sources (Table 1).

FFDAS was created using a hybrid data-assimilation approach and uses energy statistics from the International Energy Agency (IEA). Observational data from population and nightlights are used to constrain a predictive model known as the Kaya identity, which estimates the flux of $FFCO_2$ emissions from a region (Rayner et al. 2010; Asefi-Najafabady et al. 2014). Sectors in FFDAS are based on IEA emissions sectors, including emissions from commercial electricity generation, manufacture, international and domestic transport, as well as emissions from other sectors such as residential, agriculture, and fishing. It is important to note that FFDAS does not include emissions from industrial processes such as calcining limestone in the manufacture of cement and gas flaring (Rayner et al. 2010; Asefi-Najafabady et al. 2014) so the national total is slightly lower than for the other data sets (see Table 2).

The EDGAR data set was created using a hybrid approach and utilizes activity data primarily based on energy statistics produced by the IEA, but supplemented by data from British Petroleum (BP), the U.S. Geological Survey (USGS), and the United Nations (U.N.) (Marland et al. 1999; Olivier et al. 2005). EDGAR includes emissions from fuel combustion, fugitive emissions from fuel use, and industrial processes such as cement manufacture, non-energy use of lubricants/waxes, and solvent and other product uses. Emissions are calculated for the years 2000 to 2010. Emissions are distributed on a 0.1° grid using the location of energy and manufacturing facilities, road networks, shipping routes, and population density (Olivier et al. 2005; Andres et al. 2012; EDGAR 2016).

The CDIAC data set was created using a top- down approach, where nationally aggregated emissions are distributed on a regular 1° grid using a 1984 population distribution data set from the Goddard Institute of Space Studies (GISS). CDIAC relies on energy statistics produced by the United Nations (U.N.) and includes emissions from fossil fuel burning, cement manufacture, and gas flaring in oil fields (Andres et al. 1996; Marland et al. 1999; Blasing et al. 2004).

CDIAC data spans the longest time period, from 1751 to 2010 (Andres et al. 1996). The ODIAC data set was created using a hybrid approach which disaggregates national emissions estimates produced by CDIAC, but distributes them using nightlight data and the location of LPS inventoried in CARMA. Note, a previous version of ODIAC cited in the Oda and Maksyutov (2011) paper is based on BP statistics, but the current revised data are based on CDIAC statistics. ODIAC is published at the same resolution as the nightlight data, approximately 1 km, or 0.008333°, for the year 2008 (Oda and Maksyutov 2011).

Table 2 Author-documented totals compared to user calculated global and national totals

Data	Author-documented total (tonnes C)	Global-calculated total (tonnes C)	US-calculated total (tonnes C)
Vulcan	2002: 1,541,353,052.7 (USA)	N/A	1,541,353,000
ODIAC	2008: 8,468,120,000 (±2–3 %)	8,468,069,000	1,533,826,000
FFDAS	2002: 6,222,573,500	2002: 6,222,573,000	2002: 1,396,513,000
	2008: 7,617,185,500	2008: 7,617,185,000	2008: 1,380,671,000
CDIAC	2002: 6,712,000,000	2002: 6,711,645,000	2002: 1,505,423,000
	2008: 8,288,000,000	2008: 8,287,658,000	2008: 1,509,024,000
EDGAR	?	2002: 7,041,831,063	2002: 1,519,385,000
		2008: 8,654,032,698	2008: 1,499,356,000

The Vulcan data set was created using a bottom-up approach and utilizes 7 primary datasets to estimate $FFCO_2$ emissions. Vulcan uses CO_2 emissions directly when estimates are available, or CO emissions and CO_2 emissions factors as an alternative when CO_2 emissions are not reported. Vulcan estimates emissions from air transport, commercial, industrial, and residential energy use, cement manufacturing, utilities, and non-road and on-road mobile activities (Gurney et al. 2009). Emissions from LPSs are retrieved from the EPA/EIA NEI and the EPA CAMD ETS/CEMs data. County-level, road- specific emissions are distributed using a GIS road atlas with twelve road types falling under the more general road classifications of rural and urban. Non-point source emissions are available from the NEI at the county scale and are downscaled to census tracts using the total floor square footage of industrial, commercial, and residential buildings within each census tract (Gurney et al. n.d.; Gurney et al. 2009). The data are then rendered to a regular 10-km grid using area-based weighting (Gurney et al. n.d.; Gurney et al. 2009). Vulcan is often considered to be the most accurate $FFCO_2$ emissions inventory, but data on $FFCO_2$ emissions at the subnational scale are not very common, and are usually collected for very specific purposes (Andres et al. 2012). Due to limited data availability at subnational resolutions, bottom-up approaches are currently constrained to regional scales and shorter time scales. Vulcan is thus available only for the year 2002 and is limited to the spatial extent of the USA.

2.2 Methods

Data files representing the five emissions data sets were acquired from their authors and imported into ArcGIS (ESRI 2014. ArcGIS Desktop: Release 10.1. Redlands, CA: Environmental Systems Research Institute) using geographic coordinate system (GCS) WGS84. There are two primary referencing systems that assign data to a geographic location on the surface of the earth (geo-referencing systems). A spherical coordinate system represents locations of a feature in latitude and longitude in units of degrees, minutes and seconds on a spherical grid. However, rather than working with a curved surface, it can be more useful to represent the Earth's surface projected onto a flat plane. For example, paper maps are flat, and a planar representation is required to see all of Earth at one time. It is easier to perform distance and area measurements on a plane. The Earth is not a perfect sphere and representing Earth on a planar surface involves representing the Earth as a "best fit" ellipsoid. Today an international standard ellipsoid for representing the globe has been adopted, known as the World Geodetic System (WGS84) with the WGS84 datum. The decision to use GCS WGS84 for this study is based on discussions with the authors of FFDAS and ODIAC. Future comparative analysis of datasets represented using projected geo-referencing systems instead of spherical coordinate systems may prove informative. When no geo-referencing was specified for a dataset beyond spherical coordinates, the data were displayed using a WGS84 ellipsoid model. The consideration of map projection is important because the transformation from a three-dimensional world to a two-dimensional surface distorts at least one of the following characteristics: shape, area, distance or direction (Andres et al. 2012). Further, resampling techniques used to convert from units of degrees, minutes and seconds to other units found in projected coordinate systems, such as meters or feet, can affect the magnitude of re-sampled grid cells, making it important to address the geo-referencing system used when handling geographic data. Total $FFCO_2$ emissions were calculated for each inventory's largest available extent using zonal statistics in ArcGIS and compared to the author-documented totals to ensure the data were read correctly into the GIS.

All data were analyzed for the continental USA only, where the USA is defined as all 0.1° resolution grid spaces in Vulcan with non-zero values for CO_2 emissions. After the global data sets were masked to the extent of the Vulcan inventory, national level emissions were calculated using Zonal Statistics in ArcGIS. The Zonal Statistics operation in ArcGIS operates by calculating statistics of one gridded dataset based on a "zone" that is defined by another gridded dataset. The zone was defined by all Vulcan non-zero grid spaces at 0.1° resolution, and the statistic calculated was the sum of each of the five $FFCO_2$ emissions inventories within the boundary of the defined zone. Since the data were masked to the extent of the Vulcan inventory, comparisons between emissions inventories occurred only at locations where Vulcan has non-zero values. At the 0.1° resolution the maximum number of cells included in the comparison is 81,841 grid cells. There are 928 grid cells at 1° resolution. The scale of this analysis is limited to the continental USA because Vulcan, which is often considered the most detailed data set, is limited to the USA. Vulcan is also limited to the year 2002 so comparisons are completed for 2002. ODIAC, a fine resolution data set at 1 km resolution, is available only for the year 2008, so comparisons are also performed for the year 2008. All other datasets are available for both 2002 and 2008 comparisons. Due to the limited temporal extents of Vulcan (2002) and ODIAC (2008), a comparison was not made between Vulcan and ODIAC.

The data being analyzed were initially acquired at resolutions ranging from 1 km to 1°. In order to make comparisons between emissions inventories at a variety of resolutions, all data were re-sampled to the same resolution after they were masked to the Vulcan extent. ODIAC was acquired at a 1 km, or 0.008333°, resolution and was aggregated to 0.1° resolution using the *Aggregation* tool in ArcGIS with an aggregation method of sum and an aggregation factor of 12. Vulcan was analyzed using a 0.1° resolution version of Vulcan produced by the original authors of the data set. It should be noted that the 0.1° data set is provided by the Vulcan team to assist users with re-gridding and that users are encouraged to re-grid the 10 km resolution data set on their own. The authors of Vulcan assume no responsibility for re-gridding choices that do not match expectation. However, due to the distortions that are inherent with un-projecting geographic data from planar coordinates to spherical coordinates we chose not to re-grid the 10 km Vulcan data set, but to use the 0.1° data set provided by the authors of Vulcan.

EDGAR, FFDAS, and Vulcan were acquired at 0.1° resolution. EDGAR, FFDAS, ODIAC, and Vulcan were aggregated from 0.1° resolution to 0.5°, 1°, 2°, and 3° resolutions using the *Aggregate* tool in ArcGIS with an aggregation method of sum and aggregation factors of 5, 10, 20, and 30, respectively. All data were aggregated from their 0.1° resolution to each of the respective resolutions. CDIAC data were acquired at 1° resolution and were aggregated to 2° and 3° resolutions using the Aggregate tool in ArcGIS with an aggregation method of sum and aggregation factors of 2 and 3, respectively. The sum aggregation method was used for this analysis because it was necessary to preserve the national total of $FFCO_2$ emissions for each data set across different resolutions. All of the data sets except CDIAC were originally represented at 0.1° resolution or finer, so analyses were conducted at 0.1° resolution and aggregates thereof, except comparisons that included CDIAC, which were only conducted at aggregates of 1° and greater.

Comparisons between datasets were conducted using metrics found in similar analyses of $FFCO_2$ datasets (e.g. Andres et al. 1996, 2011; Marland et al. 1999; Gregg and Andres 2008; Gurney et al. 2009; Rayner et al. 2010; Oda and Maksyutov 2011; Asefi-Najafabady et al. 2014). The comparisons conducted are listed and discussed in the remainder of this text section and then illustrated in the figures and text that follow in the Results and discussion section.

The cumulative emissions curves of each emissions inventory were plotted and compared against each other at 0.1° and 1° resolutions. Emissions were arranged in ascending order and plotted on a logarithmic scale to represent the distribution of emissions magnitudes within the continental USA For visual comparison, maps of each data set were generated in ArcGIS at 0.1°, 0.5°, 1°, and 2° resolutions. Difference maps for the most and least correlated pairs of inventories at each resolution (0.1°, 0.5°, 1°, and 2°) were also generated and analyzed.

Spatial correlation coefficients compare the magnitude of one dataset at one location to the magnitude of another data set at the same location, allowing for a numeric comparison of the similarities and differences in the spatial distribution of emissions inventories. Additionally, the sum of absolute differences is a useful metric for understanding differences in the magnitude of emissions across datasets. The sum of absolute differences and spatial correlation coefficients were calculated between $FFCO_2$ emissions inventories of the same year and resolution. Spatial correlation coefficients were calculated for each data set in MATLAB®. $FFCO_2$ emissions inventories were exported from ArcGIS in geotiff format for all resolutions (0.1°, 0.5°, 1°, 2°, and 3°) and imported into MATLAB® as matrices. Matrices were reshaped into vector arrays for each resolution using the *reshape* command in MATLAB®. The MATLAB® *reshape* command reshapes a matrix by taking elements from the original matrix in a column-wise approach and rewriting them to an array of the same number of pixels. The vector array of each data set contains the same number of values as the original matrix and each matrix is reshaped according to the same column-wise approach, so each row in the new vector array represents the same location as the corresponding row in the vector array it is being compared against. After each geotiff matrix was reshaped into vector arrays, the *cov* and *corrcov* functions were used to calculate the correlation between each $FFCO_2$ emissions inventory at resolutions of 0.1°, 0.5°, 1°, 2°, and 3°. The *cov* command creates a covariance matrix between two vectors, and the *corrcov* command computes a correlation matrix corresponding to the covariance matrix generated by the *cov* command.

Correlation coefficients were calculated for each data set against the others in pairs for relevant resolutions and years. In previous studies the emissions in each dataset were scaled to the same total so that the spatial correlation coefficient was independent of magnitude (Rayner et al. 2010; Oda and Maksyutov 2011). However, because the correlation coefficient is always independent of magnitude, and both the spatial location and magnitude of emissions are important for this analysis, the authors chose not to scale the data. The sum of absolute differences (SAD) totals the absolute difference of emissions at each pixel location across all pixels and is used to measure similarity between two images. The sum of absolute differences was calculated in MATLAB® by taking the sum of the absolute differences between each set of values in the vector array.

Due to the large range of values contained within each data set, plots of the natural log of each dataset against the other in pairs (Log A versus Log B) emphasize both high and low emissions values. The natural log of the ratio of two data sets (Log of (A/B)) plotted against the mean of the two data sets (A + B)/2 compares the relative magnitude against the absolute value of the two data sets. Using this metric, values will cluster around zero when the two data points are equal, and will be symmetric about the zero axis when they differ by a similar factor (Marland et al. 1999).

Lastly, the role of LPS emissions was examined by comparing the magnitude of each $FFCO_2$ emissions inventory value at the location of the top fifty LPS emitting raster cells. LPSs inventoried in eGRID were converted from points to a raster in ArcGIS, so that each grid cell could account for emissions from multiple eGRID LPSs. The grid cells were then converted

back to points, with each point representing all eGRID sources located in a 0.1° cell. The Extract Value to Point tool in ArcGIS was used to retrieve the emissions value of each $FFCO_2$ emissions inventory at the location of the top fifty cells representing emissions from eGRID.

3 Results and discussion

3.1 Distribution of emissions

The calculated and author-documented national totals of emissions in the five studied datasets are shown in Table 2. While ODIAC is based on CDIAC statistics, the global and national totals are expected to differ within a range of 2–3 % due to corrections made in ODIAC to constrain the global total. In addition, this analysis uses a more recent version of CDIAC than the version used in the creation of ODIAC (author correspondence). After rounding errors we accept the documented totals and calculated totals to be equal. ODIAC and FFDAS have the lowest emissions estimates, while Vulcan has the highest. EDGAR's author-documented totals varied by source of documentation and whether totals were calculated across sectors or across countries.

Cumulative emissions at 0.1° (Fig. 1) show that FFDAS rises in emissions at the slowest initial rate, indicating that it attributes the majority of emissions to a small number of high-emitting cells. ODIAC follows approximately the same cumulative curve as FFDAS, but rises at a slightly faster initial rate than FFDAS, indicating that ODIAC also assigns more emissions to a small number of high-emitting cells compared to the other datasets.(Figure SF1 shows the cumulative plot for 1 degree resolution). ODIAC has only about half the number of non-zero grid spaces as the other $FFCO_2$ emissions inventories (Figure SF2). The 0.1° spatial distribution maps of ODIAC and FFDAS (Fig. 2) also illustrate this distribution of values, where cells located in less populated areas in the western U.S. do not contain nightlights and are not assigned emissions values. However, because FFDAS also uses population as a spatial proxy, it attributes emissions to more cells than ODIAC does based on nightlights alone. In both

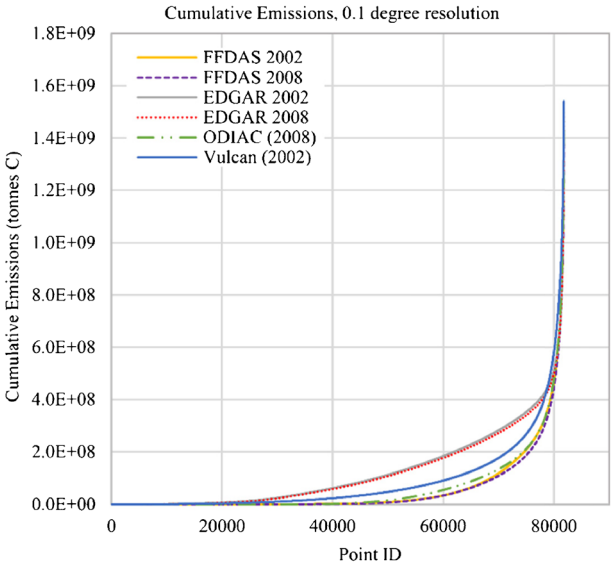

Fig. 1 Cumulative emissions curves Vulcan (2002), ODIAC (2008), FFDAS (2002 and 2008), and EDGAR (2002 and 2008) at 0.1° resolution in units of tonnes of C. Values are compared at 81,841 non-zero grid cells in the 0.1-degree resolution Vulcan dataset. *Solid lines* represent data for 2002, while *dashed lines* represent data for 2008

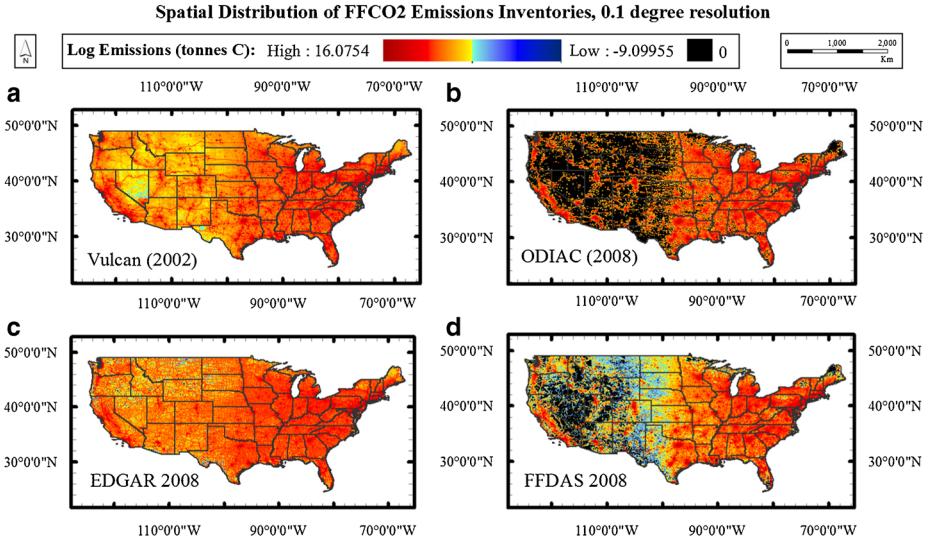

Fig. 2 Spatial distribution of the most recent years of Vulcan (**a**), ODIAC (**b**), EDGAR (**c**), and FFDAS (**d**) at 0.1° resolution. Units are expressed as the natural log of tonnes of C. Data are represented on a *log scale* in order to display the data from different inventories on the same color ramp, and to highlight both high and low emissions values. As a result of representing the data on a log scale, values of zero do not appear on the color ramp. Zero values are instead represented by the *color black*

FFDAS and ODIAC emissions are concentrated in urban areas where nightlight values and population density are high. The distribution plots for FFDAS at 0.1° resolution show discrete clustering of emissions values at lower magnitudes (Figure SF2). At the 0.1° resolution FFDAS and ODIAC attribute the same emissions value to multiple cells at lower magnitudes, creating discrete intervals of emissions without discrimination. When FFDAS and ODIAC reach low emissions values, their distribution curves drop to zero while the other data sets have long tails of decreasing values (Figures SF2 at 0.1° resolution and SF3 at 1° resolution).

The cumulative emissions curve for Vulcan at 0.1° rises initially at a faster rate than for both FFDAS and ODIAC (Fig. 1), indicating that it attributes more emissions to low and intermediate emitting cells. Vulcan, for example, assigns more emissions to the mid-west than do either FFDAS or ODIAC. The detailed, fine resolution data used to build the Vulcan $FFCO_2$ emissions inventory allows it to capture emissions in locations that cannot be predicted by nightlights, population, or large point sources alone. Vulcan's detailed methods not only distribute emissions to the west, but also to extensive transportation networks (Fig. 2a). Vulcan has the second to largest range of emissions values, just behind EDGAR.

EDGAR's cumulative emissions curve at 0.1° rises at the fastest initial rate (Fig. 1), indicating that a larger fraction of EDGAR's emissions are located in low and intermediate cells, and that its emissions values are distributed more evenly, rather than being concentrated in major urban areas. The distribution of emissions at 0.1° (Figure SF2) highlights the lower emissions values contained in EDGAR, with the lowest values in EDGAR being orders of magnitude smaller than other $FFCO_2$ emissions inventories. The more uniform distribution of emissions values in EDGAR is shown by the large range of emissions values, with the majority of cells being concentrated between 100 and 10,000 tonnes of C. EDGAR tends to estimate on the low end for major cities and other urban areas, while allocating more emissions to areas surrounding cities (Fig. 2c). EDGAR also attributes a greater amount of emissions in

the west than FFDAS, ODIAC, or Vulcan. The spatial distribution of EDGAR is unique compared to the other $FFCO_2$ emissions inventories in this analysis, but the documentation for EDGAR is not detailed enough to provide insight into how emissions are distributed and why the data exhibit such unique patterns (EDGAR 2016).

At the 1° resolution the CDIAC dataset is also available for comparison and the differences between data sets become less apparent. Figures at 1° resolution, similar to Figs 1 and 2 at 0.1°, are available in the supplemental material as Figure SF1 and in Fig. 3. Similar to 0.1° resolution, EDGAR's cumulative curve at 1° resolution rises initially at a faster rate than the other emissions inventories, indicating that it still attributes more emissions to a greater number of low and intermediate emitting cells. At 1° resolution, CDIAC, FFDAS, ODIAC, and Vulcan follow a similar path, attributing fewer emissions to low emitting cells. After approximately three-quarters of all grid cells, CDIAC's cumulative emissions curve diverges from FFDAS, ODIAC, and Vulcan, and continues to rise at a slower rate. CDIAC's cumulative emissions rise quickly in the last 100 cells, indicating that CDIAC attributes the majority of its emissions to a small number of very high-emitting cells, and that even less of CDIAC emissions are located in low emitting cells than in the other emissions inventories. The same distinct intervals of emissions values present in the distribution curve of FFDAS are visible in the distribution curve of CDIAC at 1° resolution (Figure SF3).

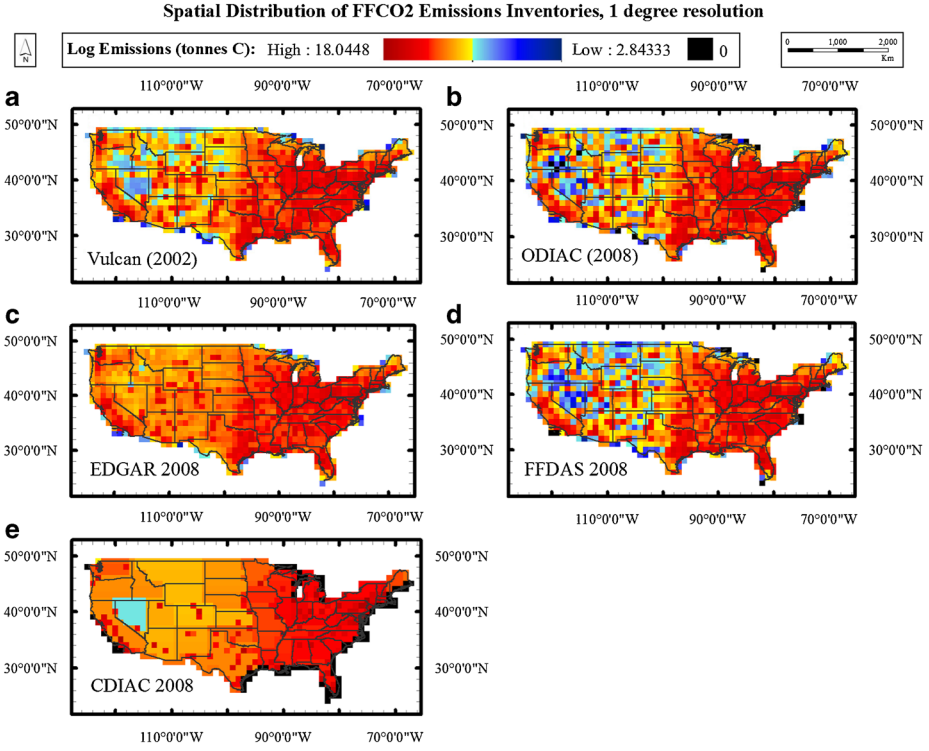

Fig. 3 Spatial distribution of the most recent years of Vulcan (**a**), ODIAC (**b**), EDGAR (**c**), FFDAS (**d**), and CDIAC (**e**) at 1° resolution. Units are expressed as the natural log of tonnes of C. Data are represented on a *log scale* in order to display the data from different inventories on the same color ramp, and to highlight both high and low emissions values. As a result of representing the data on a log scale, values of zero do not appear on the color ramp. Zero values are instead represented by the *color black*

The effect of data aggregation is shown in the spatial distribution maps at 1° resolution (Fig. 3) and 2° resolution (Figure SF4). It is important to note that because of the coarser initial resolution of CDIAC, CDIAC assigns values of zero to many cells along the US East coast and some along the West coast that Vulcan and the other emissions inventories do not (Figs 3 and Figure SF4).

While the emissions inventories become more similar at 1° resolution, each dataset holds distinct patterns of emissions that reflect their respective methods of spatial attribution. In order to further explore the relationships of emissions inventories to each other, correlation coefficients were calculated and plotted on a threshold graph to delineate the relationship between data sets at decreasing resolutions (Table 3 for 2002, Table ST1 for 2008, and Figure SF5).

Difference maps were calculated for the least and most correlated inventories for the year 2002 at 0.1° (Figure SF6) and for the years 2002 and 2008 at 1° resolutions (Figure SF7). At 0.1° resolution for the year 2002, EDGAR and FFDAS have the lowest correlation, while Vulcan and FFDAS have the highest correlation (Table 3). When FFDAS is subtracted from EDGAR it becomes apparent that EDGAR slightly overestimates relative to FFDAS in the areas surrounding major urban areas, underestimates in the majority of the urban area, and then overestimates at the very core of the city with respect to FFDAS (Figure SF6). A similar pattern appears in the difference map of EDGAR and ODIAC (not shown). Vulcan and FFDAS have large differences despite high correlation (Figure SF6).

At 1° resolution for the year 2002, CDIAC and FFDAS are the least correlated, while Vulcan and FFDAS are the most correlated (Table 3). At 1° resolution for the year 2008, CDIAC and EDGAR are the least correlated, and ODIAC and FFDAS are the most correlated (Table ST1). The difference maps between CDIAC and FFDAS for 2008 and CDIAC and EDGAR (Figure SF7) for 2008 are very similar. At 1° resolution, CDIAC tends to overestimate emissions in cells that contain major cities with respect to the other data sets. In some cases when CDIAC is compared to FFDAS, CDIAC underestimates for the cell that contains a major city, and overestimates a cell adjacent to it, such as at Houston, Cleveland, and Minneapolis. CDIAC underestimates for Baltimore, Philadelphia, and New York with respect to the other data sets due to zero-magnitude cells along the East coast (Figures SF4 and SF7). The difference maps for the most correlated inventories are similar in that Vulcan and ODIAC both differ the most from FFDAS in urban areas. However, the magnitude difference between FFDAS and Vulcan is much greater than the difference between FFDAS and ODIAC. Vulcan also has more of a tendency to either overestimate or underestimate emissions in major cities compared to FFDAS than ODIAC does when compared to FFDAS.

Table 3 Correlation coefficients and sum of absolute differences (SAD) in MtC (megatonne of Carbon) for the year 2002 at various levels of spatial aggregation from 0.1 to 3 degrees

	FFDAS vs. Vulcan		EDGAR vs. Vulcan		FFDAS vs. EDGAR 2002		EDGAR vs. CDIAC 2002		Vulcan vs. CDIAC 2002	
	Correlation	SAD	Correlation	SAD	Correlation	SAD	Correlation	SAD	Correlation	SAD
0.1	0.62	1172	0.43	1494	0.37	1404	–	–	–	–
0.5	0.91	521	0.55	663	0.89	650	–	–	–	–
1	0.95	392	0.94	468	0.94	469	0.35	1603	0.39	1650
2	0.97	292	0.97	311	0.97	320	0.71	1012	0.72	1034
3	0.98	248	0.98	271	0.98	262	0.77	839	0.79	805

In an effort to gain more insight into the difference seen between EDGAR and other emissions inventories around urban areas, profiles were taken from the SW to the NE across three major cities, Baltimore, Philadelphia, and New York. Emissions from EDGAR, FFDAS, and ODIAC at the 0.1° resolution were plotted along the transect to further explore the urban distribution of $FFCO_2$ emissions (Fig. 4). While the three data sets have good agreement over Baltimore, EDGAR overestimates at the center of both Philadelphia and New York compared to FFDAS and ODIAC. EDGAR similarly overestimates emissions at the core of other major cities, such as Los Angeles, when compared to FFDAS and ODIAC. The tendency for EDGAR to both underestimate and overestimate in comparison to FFDAS, ODIAC, and to a lesser degree, Vulcan, is demonstrated by the three prong distribution of EDGAR when plotted against FFDAS (Fig. 5a), ODIAC (Fig. 5b, c), and Vulcan (Fig. 6b). The straight lines of points in Fig. 5a and in later figures illustrate again the inability of a data set to discriminate among low values.

Many of the differences and similarities in emissions distributions presented thus far can be related to the methods and data used to create them. In order to make broader statements about the different approaches used for creating $FFCO_2$ emissions inventories, the relationships between specific datasets are further explored below. In addition to analyzing the calculated correlation coefficients, the data are plotted against each other in pairs on a log scale to evaluate differences and similarities in spatial allocation approaches such as top- down versus bottom-up, nightlights versus population density, and the treatment of large point sources.

3.1.1 Top-down vs. bottom-up

While bottom-up emissions inventories such as Vulcan are considered to be more detailed, they are time, data, and labor intensive and encounter both spatial and temporal limitations. Top-down approaches are less detailed but are less costly and are more globally consistent. In order to make comparisons between less intensive top-down approaches and more detailed bottom-up approaches, $FFCO_2$ emissions inventories for the year 2002 were compared against

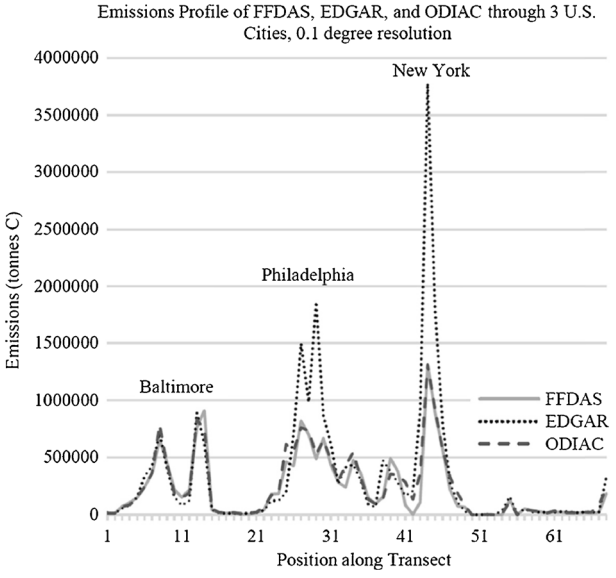

Fig. 4 Transect profile of FFCO2 emissions across three US cities: Baltimore, Philadelphia, and New York in units of tonnes of C

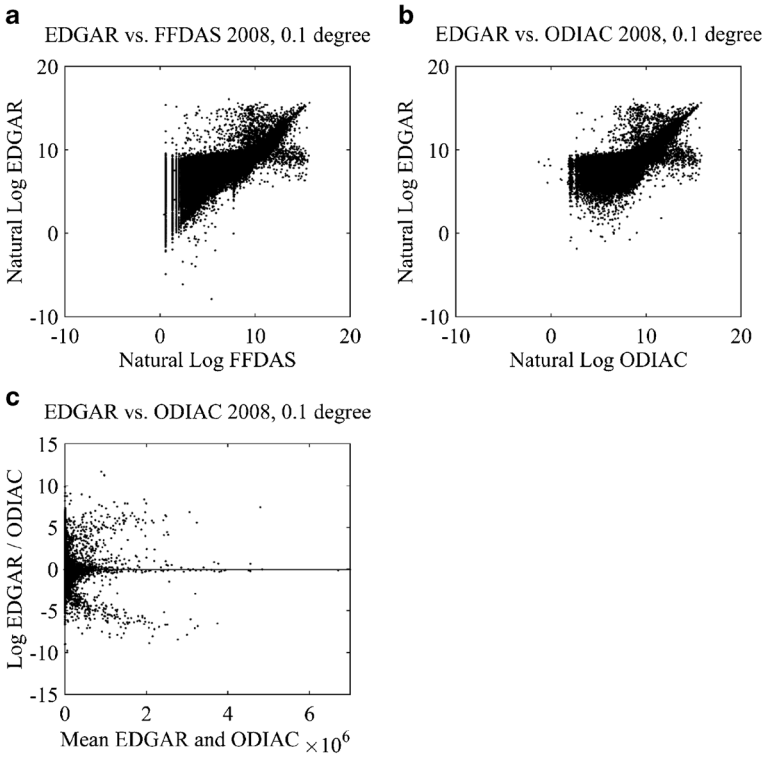

Fig. 5 The "three-prong" distribution of EDGAR, representing the tendency of EDGAR to either agree with another emissions inventory or to significantly over- or under-estimate it. Units are in log tonnes of C

Vulcan. Vulcan is considered to be the most detailed $FFCO_2$ emissions inventory, so $FFCO_2$ emissions inventories that are better correlated with Vulcan might be considered better representations of $FFCO_2$ emissions distribution. The threshold of correlation was graphed for EDGAR, FFDAS, and CDIAC against Vulcan at all comparable resolutions (Figure SF5) to show that correlation is quite high for all comparisons when the resolution is greater than about 1 degree. At 0.1° resolution, the log of FFDAS and EDGAR (2002) were plotted against the log of Vulcan (Fig. 6a, b). At 1° resolution the log of EDGAR, FFDAS, and CDIAC were plotted against the log of Vulcan (Fig. 6c–e). The figures show how much the correlation is improved in going from 0.1 to 1 degree resolution.

At the 0.1° resolution, EDGAR and Vulcan have a spatial correlation of 0.43 (Fig. 6b, Table 3). FFDAS and Vulcan are the most correlated data sets at 0.1° resolution with a correlation coefficient of 0.62 (Fig. 6a, Table 3). However, as indicated by previous results, there is a limitation in FFDAS to discriminate among emissions levels at low emissions magnitudes (Fig. 6a).

As resolution decreases to 1°, the relationship of EDGAR and FFDAS to Vulcan strengthens with correlation coefficients of 0.94 (Fig. 6d) and 0.95 (Fig. 6c), respectively (see also Table 3). At 1° resolution CDIAC is also available for analysis, showing a correlation with Vulcan of 0.39 (Fig. 6e, Table 3). As shown in the threshold graph (Figure SF5), CDIAC consistently has low correlation with Vulcan, while FFDAS and EDGAR are best correlated with Vulcan at 1° resolution.

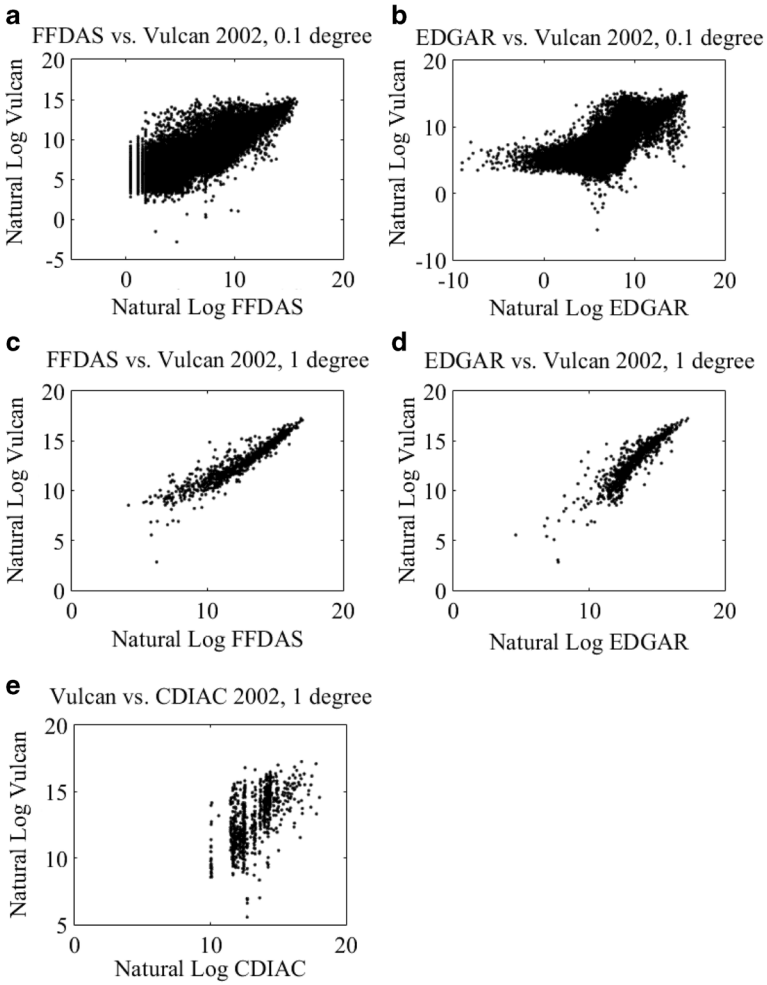

Fig. 6 Vulcan (bottom-up approach) compared to three top-down emissions inventories. Vulcan is compared to FFDAS and EDGAR at 0.1° resolution (**a**, **b**) and to FFDAS, EDGAR, and CDIAC at 1° resolution (**c**, **d**, **e**). Units are in log tonnes C

FFDAS underestimates compared to Vulcan (as seen in the displacement of the plot in Fig. 6c from a 1 to 1 line), which was expected for this analysis because FFDAS does not contain emissions from cement manufacture while Vulcan does. As a result, the correlation between FFDAS and Vulcan at 0.1° is 0.61. Asefi-Najafabady et al., (2014) calculated a correlation coefficient of 0.86 between Vulcan and FFDAS at 0.1°. However the Asefi-Najafabady (2014) analysis did not include the cement sector from Vulcan, while this analysis included emissions from all sectors (including cement) in Vulcan, because the focus was on comparing the spatial distribution of emissions inventories. The difference between these two correlation coefficients indicates that the omission of emissions from cement in FFDAS results in significant differences between FFDAS and Vulcan. When CDIAC is plotted against Vulcan on a log scale at 1°, CDIAC displays discrete intervals of un-discriminated emissions values as an artifact of the population data that CDIAC's spatial distribution is based on (Fig. 6e). As

shown in Fig. 3e, CDIAC sometimes assigns the same population density to all cells that fall within a specific state, thereby creating distinct clusters of values. At 1° resolution, Vulcan and CDIAC are poorly associated, with a spatial correlation coefficient of 0.39 (Table 3). Even at 3° resolution the correlation between Vulcan and CDIAC is only 0.79, indicating that there is a large discrepancy between the spatial distribution and magnitude of Vulcan and CDIAC data.

At a 1° resolution the differences between two top-down approaches, FFDAS and EDGAR, nearly disappear when compared to Vulcan. If $FFCO_2$ emissions inventories are going to be used at a 1° resolution, then top-down approaches differ little in their estimate and distribution of emissions compared to bottom-up approaches. If data are to be used at finer resolutions of 0.5° or 0.1°, the approach for distributing emissions begins to matter much more. However, because all of these inventories are estimates it is not possible to definitively say which approach is more accurate. Vulcan is generally considered to be the most detailed representation of emissions for the U.S. FFDAS may be considered a good global representation because it has better correlation with Vulcan for the USA. However, Vulcan and FFDAS share the same authors, so a stronger correlation is not unexpected between Vulcan and FFDAS. In addition, the relationships that FFDAS uses to distribute emissions may be most accurate for more developed countries like the USA, and the quality of FFDAS's ability to distribute emissions should not be extrapolated beyond the national extent examined in this analysis.

3.1.2 Nightlights vs. population density

Top-down $FFCO_2$ emissions inventories utilize different methods and proxy data sets for spatially disaggregating national level emissions into subnational units. ODIAC relies on nightlight data to distribute $FFCO_2$ emissions in space, while CDIAC relies on population density, for example. FFDAS utilizes both nightlight data and population in the disaggregation of $FFCO_2$ emissions. Comparisons are made between FFDAS and ODIAC, FFDAS and CDIAC, and ODIAC and CDIAC to compare the effects of using nightlight data, population, or a combination of the two in the distribution of $FFCO_2$ emissions. The relationship between ODIAC and CDIAC is especially of interest because these data sets start with the same global and national emissions totals but use different approaches to distribute them in space. The relationship between ODIAC and FFDAS is also of interest because both datasets use nightlight data to distribute emissions.

ODIAC and CDIAC use the same national emissions estimates but distribute national level emissions using different approaches. ODIAC uses nightlights data and the location of large point sources as inventoried in CARMA to distribute emissions on a 1 km grid, while CDIAC uses a 1984 population density map to distribute emissions on a 1° grid. The spatial relationship between ODIAC and CDIAC emissions estimates is poor, with a correlation coefficient of 0.38 at 1° resolution and 0.76 at 3° resolution (Fig. 7, Table ST1).

At 0.1° resolution FFDAS and ODIAC have the best fit at cells with higher emissions magnitudes (Fig. 8a). The greatest agreement between ODIAC and FFDAS at higher emissions values reflects the tendency of both FFDAS and ODIAC to concentrate the majority of their emissions to high-emitting cells in urban areas. The distribution of ODIAC and FFDAS become much more similar at the 1° resolution and above than at finer resolutions, trending at the same values throughout the distribution (Fig. 8). As resolution decreases the fit between FFDAS and ODIAC increases. At 0.1° resolution the correlation of FFDAS and ODIAC is 0.92, which jumps to 0.98 at 0.5° resolution, and 0.99 at 1° resolution and lower (Fig. 8, Table ST1). As resolution decreases to 2° it becomes apparent that FFDAS has slightly lower

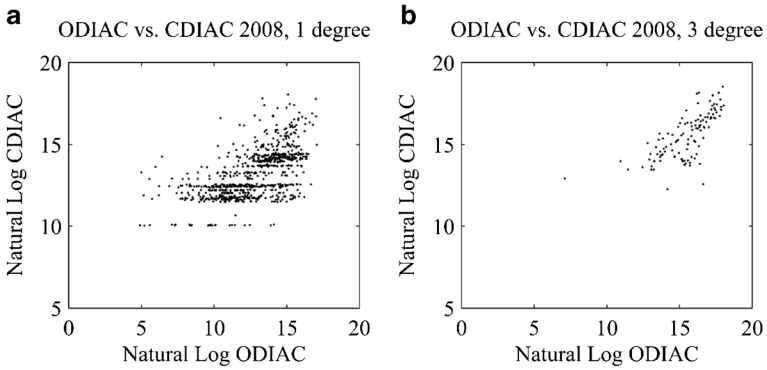

Fig. 7 Log of ODIAC versus the log of CDIAC at 1° and 3° resolutions. ODIAC and CDIAC use the same global and national totals but distribute emissions at the subnational scale using different spatial proxies. Units are in log tonnes C

values than ODIAC. When the absolute value of FFDAS and ODIAC are compared to their relative value, the emissions magnitudes cluster just below zero, reflecting the lower national total of emissions estimated by FFDAS.

As previously indicated, FFDAS and CDIAC display discrete clustering of emissions values (see, for examples Figs 6a, e and Figure SF2 and Figure SF3). However, this pattern is also seen

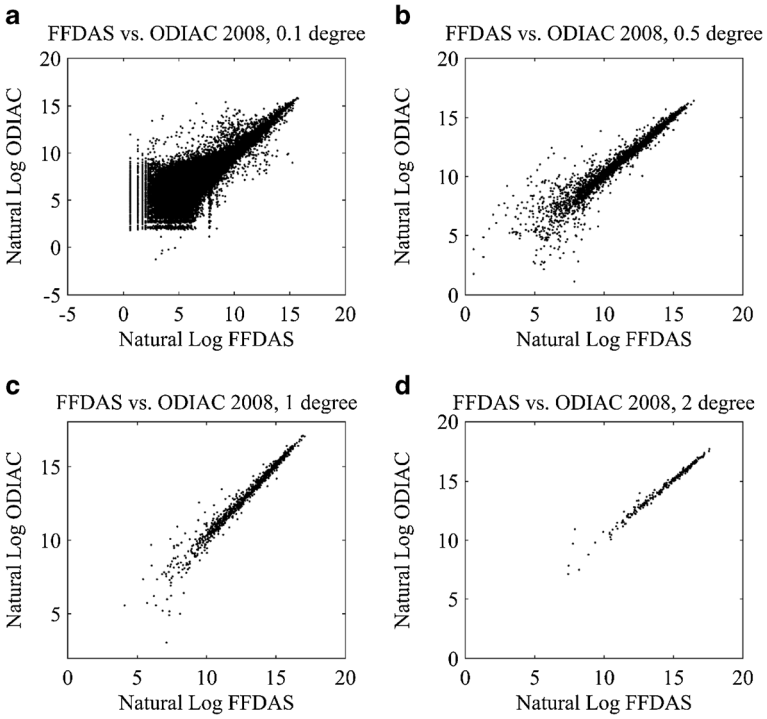

Fig. 8 Log of FFDAS versus the log of ODIAC at 0.1°, 0.5°, 1° and 2° resolutions. ODIAC and FFDAS use different global and national emissions totals, but both use nightlight data as a spatial proxy to distribute emissions subnationally. FFDAS also uses population as a spatial proxy, while ODIAC does not. Units are in log tonnes C

in ODIAC. When CDIAC (Fig. 7a) and FFDAS (Fig. 8a) are plotted against ODIAC on a log scale at 0.1 degree resolution, clustering of emissions values is visible in all three emissions inventories, signifying that this pattern is related to the use of nightlights and/or population in the distribution of emissions. This suggests that the use of nightlights data and population do not provide a good discriminator of emissions at lower magnitudes. The clustering seen in CDIAC only occurs because of the coarse resolution population data it is based on. Since FFDAS and ODIAC both display discrete values of emissions at low values, and both use nightlight data, this pattern may be reflective of a limitation in nightlight data also.

After FFDAS and ODIAC are aggregated to 1° resolution the range of emissions values decreases and the clustering of data values become less distinct. At 3° resolution the discrete intervals in CDIAC disappear when plotted against ODIAC (Fig. 7b). At 1° resolution, the limitation in nightlight data at low values may not be significant. However the difference between FFDAS and CDIAC, and ODIAC and CDIAC is poor at all resolutions, indicating that the 1° resolution population data set that CDIAC uses does not produce good agreement with other emissions inventories.

3.1.3 Treatment of large point sources

While large differences between spatially explicit $FFCO_2$ emissions inventories may be the result using of different disaggregation methods, it appears likely that the largest difference between the data sets occurs due to the treatment of large point sources. If large point source locations are included in the distribution of $FFCO_2$ emissions, nearly half the national total of emissions will be accurately allocated in space. When high emissions values are not directly related to an urban area, they are likely associated with emissions from large point sources.

Emissions from the top 50 CO_2 emitting grid spaces from the eGRID database for the year 2009 are plotted against values of emissions from each $FFCO_2$ emissions inventory at the corresponding location at 0.1° (Figure SF8) and 1° (Fig. 9) resolutions. This is indicative of the extent to which the various inventories capture emissions from LPSs located within a particular grid cell at each respective resolution, in spite of the mismatch in inventory years (the Vulcan data are for 2002 and the other inventory data are for 2008).

It should be noted that CARMA is a more commonly used LPS inventory for global $FFCO_2$ emissions inventories than eGRID because it is global while eGRID is for the U.S. only. At 0.1° resolution (Figure SF8) FFDAS and ODIAC have the best agreement with emissions from large point sources, reflecting the incorporation of large point sources into these emissions inventories. At 0.1° resolution Vulcan and EDGAR apparently have poor agreement with emissions from LPS's. This pattern is unexpected as Vulcan places all point values in the grid cell occupied by their geocoded location, and the emissions mismatch may be partly attributed to the small temporal mismatch. Also, the Vulcan inventory uses neither eGRID nor CARMA data to estimate emissions from electrical generation, but relies on the EPA ETS/CEMs data, which is one input into eGRID, as well as the NEI point source emissions data. However, CEMs measured emissions do not always agree with emissions calculated using material balance. In addition to using different LPS data than eGRID, Vulcan is developed for the year 2002, while the eGRID data in this analysis estimates emissions for the year 2009. EDGAR reports the use of point source data in their methodology, although no reference to a specific LPS inventory is given.

At 1° resolution (Fig. 9) all $FFCO_2$ emissions inventories have values consistently equal to or higher than the large point sources located in that grid cell except for CDIAC. At the 1°

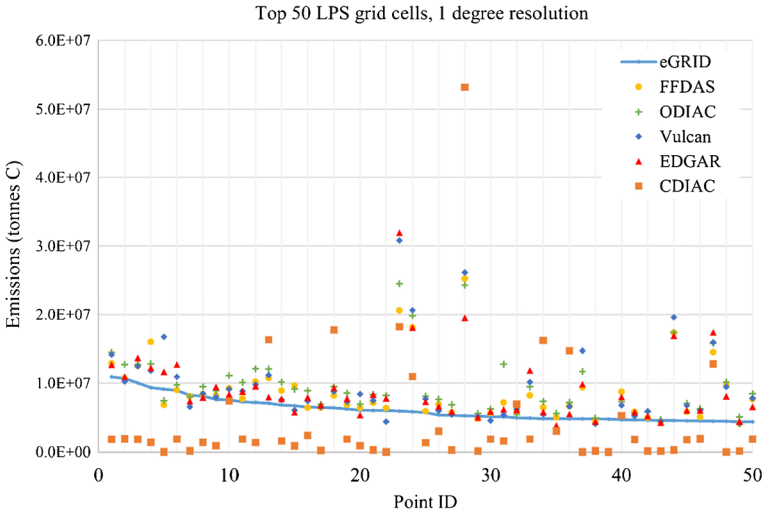

Fig. 9 The emissions values of FFDAS, ODIAC, Vulcan, EDGAR, and CDIAC corresponding to the top 50 electricity-generating LPS emitting grid cells from eGRID. Point sources from eGRID (2009) were converted to a 1° grid and the top 50 emitting cells were compared to emissions values of each $FFCO_2$ emissions inventory at 1° resolution. The Point ID is a unique ID for each value plotted

resolution CDIAC has consistently lower values than other emissions inventories at cells corresponding to the location of the top 50 LPS emitting cells for 2009 (Fig. 9). In the instances where CDIAC does exceed emissions magnitudes that can be attributed to LPS, the cells are associated with major urban areas.

3.2 Geo-referencing

Few $FFCO_2$ emissions inventories explicitly describe the geo-referencing system used to analyze and represent the data. The selection of a map projection commonly relies on the projection of the data set used to delineate the borders within which the $FFCO_2$ emissions will be distributed (Andres et al. 2012). Traditionally, $FFCO_2$ emissions inventories are represented on a regular 0.1° grid using spherical coordinates. All data sets analyzed use spherical coordinates, represented in GCS WGS84. Andres et al. (2012) address the importance of map projections, recognizing that the conversion from a three-dimensional world to a two-dimensional surface can distort shape, area, distance, or direction. Mathematical equations exist to transform data from one projection to another, but the distortion caused by conversions between planar coordinate systems and three-dimensional spherical representations of the earth are important considerations that should be addressed in future analyses. Map projections are especially important when $FFCO_2$ emissions inventories are going to be used in a model with a specific, and possibly different, map projection built into it (Andres et al. 2012).

3.3 Scale issues

The power of a geographic information system (GIS) lies in its ability to transform, analyze and manipulate geographic data; however, all of these abilities rely on measurements of scale (Goodchild 2011). Primary questions considered in geographic studies include what are the

appropriate extent and resolution at which to examine a specific geographic phenomenon (Cao and Lam 1997). The concept of ecological fallacy describes the practice of attributing characteristics of data from a large scale to smaller scales (Cao and Lam 1997; Goodchild 2001, 2011). Global $FFCO_2$ emissions totals are not used to characterize national level $FFCO_2$ emissions totals since total CO_2 emissions at the national scale do not agree with total emissions at the global scale. Therefore, attributing national level emissions to subnational scales may represent a form of ecological fallacy, in which data at a larger scale are erroneously attributed to a smaller scale. In addition, while there are proxies used in the disaggregation of $FFCO_2$ emissions inventories, the relationships between proxies are scale-dependent. For example, population density is not a good proxy for $FFCO_2$ at fine scales (Andres et al. 2011). If top-down approaches incorporate errors of ecological fallacy and scale dependence, bottom-up approaches may be a stronger approach for estimating subnational $FFCO_2$ emissions. Data are often only available at scales that are too coarse for modeling a given process at an appropriate scale. However, because the Vulcan inventory was acquired at such fine scales it is possible to measure the impacts of using data that are too coarse for their purpose. While some of the proxies used to disaggregate emissions are scale dependent, future simulations may address more accurate and complex methods for disaggregation.

When masking global $FFCO_2$ emissions inventories to the Vulcan extent, national level emissions from one inventory may not be captured if that emissions inventory uses a different boundary definition, making comparisons of U.S. national totals across $FFCO_2$ emissions inventories using this methodology imprecise. One example of differing boundaries is shown in Fig. 3e and SF4e, where CDIAC contains zero-magnitude cells along the border where other emissions inventories assign non-zero values. In general, the effects of border issues can be seen in lower emissions values along the outward edge of all emissions inventories at 0.5° (not shown), 1° (Fig. 3), and 2° (Figure SF4) resolutions. There are also higher standard deviations along boundaries in this analysis, indicating poor agreement along the edges of clipped data (Figure SF9e and SF9f).

Additional issues arise when masking data at the global scale to a national extent, in determining whether it is more appropriate to first aggregate the data to the global scale, and then mask it to the national scale, or to first mask the data and then aggregate. If the data are first aggregated and then masked to the national extent, the coarser resolutions capture a higher magnitude of emissions from outside the national boundary, causing emissions at the national level to increase with resolution. Aggregating the global data sets after they have been masked produces consistent national totals but misleadingly excludes data values that may be relevant, depending on the national boundary definition for that data set. In this analysis, smaller magnitudes and higher standard deviations are seen in cells along the Canada and Mexico borders, where emissions at the global scale were clipped to the national extent of Vulcan.

Across each comparison made in this study, aggregating the $FFCO_2$ emissions inventories to coarser resolutions produced better correlation between the datasets. However, while the correlations have increased, previous studies indicate that improved correlations at aggregated resolutions tend to be offset by the loss of degrees of freedom, making relationships between aggregated data no more significant (Goodchild 2011).

3.4 Uncertainty across emissions inventories

While comparing $FFCO_2$ emissions inventories against the spatial distributions, magnitudes, and inputs of other $FFCO_2$ emissions inventories lends great insight into the effects of methods and

data on $FFCO_2$ emissions estimates, it does not fully capture the uncertainty that arises from the variability of $FFCO_2$ emissions estimates. Attempts have been made to estimate the uncertainty associated with a single $FFCO_2$ emissions inventory, for example, in the case of FFDAS, where posterior uncertainty is calculated using Monte Carlo simulations (Asefi-Najafabady et al. 2014). However, no multi-model uncertainty estimate has been previously estimated.

In a first attempt to estimate uncertainty across $FFCO_2$ emissions inventories, maps of the average emissions for the USA for the year 2002 were generated at 0.1° and 1° resolution (Figures SF9a and SF9b). Averages for the year 2002 include emissions from EDGAR, FFDAS, and Vulcan at 0.1° resolution and EDGAR, FFDAS, Vulcan, and CDIAC at 1° resolution. The standard deviation across datasets was also mapped for the year 2002 at 0.1° and 1° resolution (Figure SF9c and SF9d). The standard deviation was divided by the mean to produce a map of the coefficient of variation, which compares the relative magnitude to the variance at each grid cell (Figure SF9e and SF9f).

The spatial distribution of emissions is not notably different at 0.1° or 1° resolutions. Emissions magnitudes have the highest standard deviations where the most emissions are located, i.e. in major urban areas (Figure SF9c and SF9d). Areas which have the highest emissions are the most important locations for emissions inventories to agree on, as these areas have greater impact on $FFCO_2$ emissions and have greater policy implications for the monitoring and reduction of $FFCO_2$ emissions. However, the coefficient of variance indicates that the greatest relative disagreement between emissions inventories occurs at lower emissions values.

4 Conclusions

Each of the approaches compared in this paper is known to have both strengths and weaknesses. Bottom-up approaches are more detailed, but are also more time, data, and labor intensive. Top-down approaches are less resource intensive but are not as detailed and rely on fuzzy relationships that change with scale (Rayner et al. 2010). For example, disaggregation by nightlights relies on a relationship between nightlight data and population, and correlations between population and CO_2 emissions that are not exact (Rayner et al. 2010; Oda and Maksyutov 2011). Hybrid approaches which use increasingly finer scale data at the building, road, or lower resolution, in conjunction with the more traditional top-down proxies such as nightlights and population are particularly well suited for the development of detailed global-scale, spatially-explicit emissions inventories.

At coarse resolutions (1° and higher) distinctive differences in the patterns are hard to detect, but at finer resolutions distinct patterns and relationships between datasets become apparent as functions of inputs and methods used in their creation. The results presented in this analysis reveal that the primary differences between data sets are related to differences in top-down and bottom-up approaches, the use of nightlights versus population density as proxy in the spatial distribution of emissions, and the treatment of large point sources (LPS) in the emissions inventory. The best fit between two published $FFCO_2$ emissions inventories occurs between FFDAS and ODIAC, both of which rely heavily on nightlights data. CDIAC shows the least correlation to the other $FFCO_2$ emissions inventories. As a function of CDIAC's coarse resolution it is unable to fully capture the fine distribution of emissions that occurs subnationally. Additionally, the lack of consideration for emissions from LPS in CDIAC increases differences between CDIAC and the other four emissions inventories. On the other hand, the reliance of CDIAC on only population data permits a long, consistent time series

over periods where satellite data and data on LPS are not available. The use of point source data is observed as being very important when good point source data are available.

A challenge is that although characterization of LPS appears to be very important for accurate description of the spatial distribution of emissions, especially at high spatial resolution, the available data are very limited temporally and still laden with considerable uncertainty. In the USA, the eGRID data set covers only 10 of the years in the 1996–2010 interval Woodard et al. 2014 show also that even in the USA we have to deal with both locational and magnitude uncertainty. This analysis shows that at spatial resolutions of 1° or greater different analytical approaches yield highly correlated results, but at higher spatial resolution, estimates of emissions will carry high uncertainty until LPS can be well characterized.

The unique patterns presented in EDGAR should be further analyzed after more detailed metadata on the methodologies can be obtained. In order for these $FFCO_2$ emission inventories to be used to their fullest potential more detailed metadata should be published and maintained for each data set. Vulcan has the most comprehensive metadata documentation, allowing for a better understanding of how the inputs and approaches influence the spatial distribution of emissions compared to other inventories. Future research should also investigate the strengths and weaknesses of nightlight data in the distribution of emissions related to specific sectors. For example, nightlights data may be more strongly related to residential or commercial emissions versus transportation and industrial sources. The nightlights data are not able to provide detailed discrimination at low emissions levels. An expanded analysis into the differences in urban emissions profiles for major cities could also lend insight into what is happening in these very important centers of $FFCO_2$ emissions.

The differences among these five spatial $FFCO_2$ emissions inventories are visible both graphically and numerically. While there is better correlation between the data sets at coarser resolutions, there is also a considerable amount of information loss in aggregating the data. The comparison of these five $FFCO_2$ emissions inventories across multiple resolutions brings into question whether or not coarser resolution emissions inventories are more accurate because they have better agreement with each other, or if the loss of information makes the use of aggregated data more uncertain. The question also remains whether the highest resolution data are more a reflection of the emissions proxy than of actual emissions and whether different proxies are perhaps more appropriate for particular sectors or fuel types. Data users should be aware of the strong relationship between uncertainty and scale.

While atmospheric concentrations of carbon and carbon emissions from fossil fuel combustion are considered the least uncertain values in the global carbon cycle, there is still a lot of uncertainty associated with our understanding of $FFCO_2$ emissions. Uncertainties in the global carbon cycle limit our ability to effectively, remotely measure, monitor, or verify where, when, and how much carbon is being emitted from fossil fuel sources. However, by developing a cohesive carbon monitoring system that seeks to develop best practice methodologies for quantifying carbon, uncertainties in the carbon cycle can eventually be reduced. In addition, such a carbon monitoring system will allow for the monitoring and verification of international agreements, which seek to limit carbon emissions globally (see, for examples, JPL 2015 and references therein; and Bovensmann et al. 2010).

Ultimately, the decision to use top-down versus bottom-up approaches depends on what the data are going to be used for. Does the purpose justify the cost required for more detailed analysis or are lower resolution data sufficient? In either case the importance of focusing on emissions from large point sources is apparent. All of the data sets studied here agree that emissions from the USA are concentrated in a small fraction of the total number of grid spaces.

Emissions from large point sources comprise a large fraction of the total and they are the fraction most amenable to independent monitoring and verification from remote measurements. LPS are also a common target of mitigation initiatives. Urban initiatives such as the C40 project and the Compact of Mayors are driving carbon and greenhouse gas mitigation in urban areas, with most mitigation action focused on emissions from urban buildings (C40 Cities 2016; Compact of Mayors 2016). The Stockholm institute found that CO_2 mitigation across all urban actions (building, transport, and waste management) combined could reduce global emissions by 3.7 Gt CO_2e in 2030 and 8.0 Gt CO_2e in 2050 (Erickson and Tempest 2014). Large point sources and urban concentrations of emissions appear to provide useful targets for monitoring mitigation progress or emission commitments. Fine resolution emissions inventories can guide the development of mitigation action plans in urban areas and verify their success. Different approaches to emissions inventories may be more suitable where long time series are needed, or where baseline LPS data are missing or unreliable.

Acknowledgments The data used or referenced in this paper are available at the following locations:
CDIAC gridded $FFCO_2$ emissions distribution can be found at http://cdiac.ornl.gov.
EDGAR gridded $FFCO_2$ emissions distribution can be found at http://edgar.jrc.ec.europa.eu.
FFDAS gridded $FFCO_2$ emissions distribution can be requested at http://gurney.faculty.asu.edu/research/ffdas.php.
ODIAC gridded $FFCO_2$ emissions distribution can be requested at http://odiac.org.
Vulcan data can be found at http://vulcan.project.asu.edu/research.php.
CARMA data can be found at http://carma.org/.
eGRID LPS data can be found at http://www.epa.gov/cleanenergy/energy-resources/egrid/index.html.
IEA statistics can be found at http://www.iea.org/statistics/.
DOE/EIA statistics can be found at: http://www.eia.gov/.
U.N. statistics can be found at: http://unstats.un.org/unsd/energy/.
BP statistics can be found at http://www.bp.com/en/global/corporate/about-bp/energy-economics.html.
Funding for this research comes from the Carbon Monitoring System Program (NNH11ZDA001N-CMS) of the U.S. National Aeronautics and Space Administration.
We thank Dr. Bob Andres, Dr. Kevin Gurney, and Dr. Tom Oda for their technical and logistical support and assistance in navigating these data. The comments of two anonymous reviewers have helped us to tighten up the text and improve its readability.

References

Andres RJ, Marland G, Fung I, Matthews E (1996) A 1° by 1° distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950–1990. Global Biogeochem Cycles 10:419–429. doi:10.1029/96GB01523

Andres RJ, Gregg JS, Losey L, Marland G, Boden TA (2011) Monthly, global emissions of carbon dioxide from fossil fuel consumption. Tellus B 63:309–327. doi:10.1111/j.1600-0889.2011.00530.x

Andres RJ, Boden TA, Bréon F et al (2012) A synthesis of carbon dioxide emissions from fossil-fuel combustion. Biogeosci Discuss 9:1299–1376. doi:10.5194/bgd-9-1299-2012

Asefi-Najafabady S, Rayner PJ, Gurney KR et al (2014) A multiyear, global gridded fossil fuel CO_2 emission data product: evaluation and analysis of results: global fossil fuel CO_2 emissions. J Geophys Res Atmos 119: 10213–10231. doi:10.1002/2013JD021296

Blasing TJ, Marland G, Broniak C (2004) Estimates of annual fossil-fuel CO_2 emitted for each state in the U.S.A. and the District of Columbia for each year from 1960 through 2001. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN, U.S.A., DOI:10.3334/CDIAC/00003

Bovensmann H, Buchwitz M, Burrows JP, Reuter M, Krings T, Gerilowski K, Schneising O, Heymann J, Tretner A, Erzinger J (2010) A remote sensing technique for global monitoring of power plant CO_2 emissions from space and related applications. Atmos Meas Tech 3:781–811

C40 Cities (2016) C40 Cities Climate Leadership Group. http://www.C40.org. Cited Feb 15 2016.

Cao C, Lam NS (1997) Understanding the scale and resolution effects in remote sensing. In: Quattrochi DA, Goodchild MF (eds) Scale in remote sensing and GIS. CRC/Lewis Publishers, Boca Raton

Compact of Mayors (2016) UN Secretary-General's Special Envoy for Cities and Climate Change. http://www.compactofmayors.org. Cited Feb 15 2016.

Dai J, Rocke DM (2000) A GIS-based approach to spatial allocation of area source solvent emissions. Environ Model Softw 15:293–302. doi:10.1016/S1364-8152(00)00004-9

EDGAR (2016) EDGARv4 methodology. Available at http://edgar.jrc.ec.europa.eu/methodology.php. Cited 15 Jan 2016

Erickson P, Tempest K (2014) Advancing climate ambition: How city-scale actions can contribute to global climate goals. In: Davis M (ed) SEI Working Paper No. 2014-06. Stockholm Environment Institute

Goodchild MF (2001) Models of scale and scales of modelling. In: Tate NJ, Atkinson PM (eds) Modelling scale in geographical information science. Wiley, West Sussex

Goodchild MF (2011) Scale in GIS: an overview. Geomorphology 130:5–9. doi:10.1016/j.geomorph.2010.10.004

Gregg JS, Andres RJ (2008) A method for estimating the temporal and spatial patterns of carbon dioxide emissions from national fossil-fuel consumption. Tellus B 60:1–10. doi:10.1111/j.1600--0889.2007.00319.x

Gurney KR, Mendoza DL, Zhou Y et al (2009) High resolution fossil fuel combustion CO_2 emission fluxes for the United States. Environ Sci Technol 43:5535–5541. doi:10.1021/es900806c

Gurney KR, Mendoza D, Geethakumar S et al (n.d.) Vulcan science methods documentation 2.0. Available at http://vulcan.project.asu.edu/pdf/Vulcan.documentation.v2.0.online.pdf. Cited 15 Feb 2013

Hartmann DL, Tank AMGT, Rusticucci M et al (2013) Observations: atmosphere and surface. In: Stocker TF, Qin D, Plattner GK et al (eds) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge and New York

JPL (2015) NASA megacities carbon project, Jet Propulsion Laboratory, California Institute of Technology, https://megacities.jpl.nasa.gov/portal

Marland G, Brenkert A, Olivier J (1999) CO_2 from fossil fuel burning: a comparison of ORNL and EDGAR estimates of national emissions. Environ Sci Policy 2:265–273. doi:10.1016/S1462-9011(99)00018-0

Oda T, Maksyutov S (2011) A very high-resolution (1 km × 1 km) global fossil fuel CO_2 emission inventory derived using a point source database and satellite observations of nighttime lights. Atmos Chem Phys 11:43–556. doi:10.5194/acp-11-543-2011

Olivier JG, Van Aardenne JA, Dentener FJ et al (2005) Recent trends in global greenhouse gas emissions: regional trends 1970–2000 and spatial distribution of key sources in 2000. Environ Sci 2:81–99. doi:10.1080/15693430500400345

Peylin P, Houweling S, Krol MC et al (2011) Importance of fossil fuel emission uncertainties over Europe for CO_2 modeling: model intercomparison. Atmos Chem Phys 11:6607–6622. doi:10.5194/acp-11-6607-2011

Rayner PJ, Raupach MR, Paget M et al (2010) A new global gridded data set of CO_2 emissions from fossil fuel combustion: methodology and evaluation. J Geophys Res 115(D19):1–11. doi:10.1029/2009JD013439

Singer AM, Branham M, Hutchins MG et al (2014) The role of CO_2 emissions from large point sources in emissions totals, responsibility, and policy. Environ Sci Policy 4:190–200. doi:10.1016/j.envsci.2014.08.00

Vogel FR, Thiruchittampalam B, Theloke J, Kretschmer R, Gerbig C, Hammer S, Levin I (2013) Can we evaluate a fine-grained emission model using high-resolution atmospheric transport modelling and regional fossil fuel CO_2 observations? Tellus B 2013(65):18681. doi:10.3402/tellusb.v65i).18681

Wang R, Tao S, Ciais P et al (2012) High resolution mapping of combustion processes and implications for CO_2 emissions. Atmos Chem Phys Discuss 12:21211–21239. doi:10.5194/acpd-12-21211-2012

Wheeler D, Ummel, K (2008) Calculating CARMA: global estimation of CO_2 emissions from the power sector, Working Paper 145. Center for Global Development, 1–37

Woodard D, Branham M, Buckingham G et al (2014) A spatial uncertainty metric for anthropogenic CO_2 emissions. Greenh Gas Meas Manag 4:139–160. doi:10.1080/20430779.2014.1000793