Bias present in US federal agency power plant CO_2 emissions data and implications for the US clean power plan

2016 Environ. Res. Lett. 11 064005

(http://iopscience.iop.org/1748-9326/11/6/064005)

Environmental Research Letters

OPEN ACCESS

RECEIVED
17 August 2015

REVISED
30 November 2015

ACCEPTED FOR PUBLICATION
5 May 2016

PUBLISHED
1 June 2016

LETTER

Bias present in US federal agency power plant CO_2 emissions data and implications for the US clean power plan

K R Gurney[1,2], J Huang[1] and K Coltin[3]

[1] School of Life Sciences, Arizona State University, Tempe, AZ 84287, USA
[2] Global Institute of Sustainability, Arizona State University, Tempe, AZ 85287, USA
[3] School of Statistics and Mathematical Sciences, Arizona State University, Tempe, AZ 85287, USA

E-mail: kevin.gurney@asu.edu

Keywords: climate change, climate policy, carbon dioxide, power plants

Supplementary material for this article is available online

Abstract

Power plants constitute roughly 40% of carbon dioxide (CO_2) emissions in the United States. Climate change science, air pollution regulation, and potential carbon trading policies rely on accurate, unbiased quantification of these large point sources. Two US federal agencies—the Department of Energy and the Environmental Protection Agency—tabulate the emissions from US power plants using two different methodological approaches. We have analyzed those two data sets and have found that when averaged over all US facilities, the median percentage difference is less than 3%. However, this small difference masks large, non-Gaussian, positive and negative differences at individual facilities. For example, over the 2001–2009 time period, nearly one-half of the facilities have monthly emission differences that exceed roughly $\pm 6\%$ and one-fifth exceed roughly $\pm 13\%$. It is currently not possible to assess whether one, or both, of the datasets examined here are responsible for the emissions difference. Differences this large at the individual facility level raise concerns regarding the operationalization of policy within the United States such as the recently announced Clean Power Plan. This policy relies on the achievement of state-level CO_2 emission rate targets. When examined at the state-level we find that one-third of the states have differences that exceed 10% of their assigned reduction amount. Such levels of uncertainty raise concerns about the ability of individual states to accurately quantify emission rates in order to meet the regulatory targets.

1. Introduction

Emission of carbon dioxide (CO_2) from the combustion of fossil fuels remains the largest net annual flux of greenhouse gases to the Earth's atmosphere [1]. Quantification of fossil fuel CO_2 emissions began with global and national-scale accounting but recent needs have placed more emphasis on quantification at smaller space and time scales in addition to detail regarding the emitting process [2–4]. This need is partly driven by the increasing density of atmospheric CO_2 measurements from both ground-based and remote sensing platforms [5, 6]. Utilizing such measurements within a carbon monitoring system (CMS) requires improved quantification of emissions, their uncertainties and disaggregation in space, time and by function [7, 8]. It is anticipated that a mature CMS can act to verify emissions at varying domains from city, to national and global scales. In addition to verification and improved understanding of complete carbon budgets, high-resolution emissions quantification can offer much more precise, reliable information on mitigation options and their reduction potential [9, 10]. Central to all of this research and applied policy needs is a reliable estimate and understanding of fossil fuel CO_2 emissions uncertainty. However, because much of the data used to construct high-resolution bottom-up emissions data products contain limited information on measurement or estimation procedures, uncertainty quantification remains challenging.

The United States collects considerable information regarding fuel consumption, economic activity and pollution statistics, offering a unique opportunity to understand fossil fuel CO_2 emissions uncertainty. Of all the emitting activities, electricity production represents the single largest CO_2 emitting sector in the US, accounting for roughly 40% of national emissions [11]. Electricity production is also the sector for which particularly detailed data are collected and archived but for which questions have been raised challenging the somewhat traditional assumption that it may be the most accurately estimated emitting sector. Hence, it is a logical choice for exploration of fossil fuel CO_2 emissions uncertainty and implications of uncertainty for greenhouse gas mitigation policies.

On 3 August, 2015, the need to understand and quantify CO_2 emissions and uncertainty from the production of electricity intensified. On that date, the US Environmental Protection Agency announced regulation of power plants burning fossil fuels [12]. The proposed regulation establishes numerical targets to be met by 2030 for each US state's power plants in the form of a state average emission rate. The emission rate is the quantity of CO_2 emitted per unit of electricity produced (e.g. lbs CO_2/MWhr). Because the overall goal of this proposed rule is to reduce the atmospheric burden of greenhouse gases, the accuracy of the emitted CO_2 amount within the calculation of the emission rate, is a critical element in establishing, implementing and verifying the emission rate goals. Estimation methods that can provide an unbiased emission quantity with a known level of accuracy and precision are needed. Furthermore, because each state will propose and implement the means by which they meet the EPA proposed emission rate targets, these methods must be consistent across all US power plants and transparent to the public.

An exploration of $FFCO_2$ emissions from electricity production facilities in the US can be accomplished through close examination of two datasets on power plant characteristics. The Department of Energy's Energy Information Administration (EIA) and the Environmental Protection Agency's Clean Air Markets Division (CAMD) maintain independent data collection efforts, which can be used to quantify fossil fuel CO_2 emissions at all large electricity producing facilities in the US. Because the CAMD and EIA datasets are generated by two different Federal US agencies with different mandates, the data reflect different goals and collection methods. The EPA's data collection effort is focused on establishing regulatory compliance of SO_2 and NO_x emissions and primarily uses stack monitoring. The EIA, by contrast, is focused on maintaining statistics on fuel consumption and electricity production and hence, relies primarily on fuel calculation procedures to estimate emissions.

Recent research explored the differences between the EPA's emissions data and that supplied by the EIA [13]. The annual relative difference for the total of all paired fossil fuel-burning facilities in the year 2004 was 2.5% (EPA > EIA). However, the mean individual relative difference (IRD) for all paired fossil fuel-burning facilities was 0.7% (EPA > EIA) and the mean individual 'unsigned' difference for all paired facilities was 18.3%. This suggests that the small total and average annual differences are caused by cancellation of large positive and negative individual paired differences.

A more narrowly focused study compared the 2009 EIA and CAMD CO_2 emissions at 210 coal-fired power plants, a subset of the total capacity in the US, and concluded that annual emissions from the EIA calculations were more accurate than the measured values contained within the CAMD data [14]. Though important in confirming that these two datasets have numerical differences at the facility level, a number of questions remain regarding the reliability of this analysis [15]. Indeed, the peer-reviewed discussion that followed this paper questioned the veracity of the analysis and whether or not the conclusions were possible given the limitations in the data.

In spite of these important contributions, a number of questions remain regarding the differences between these two datasets. Firstly, key attributes need to be tested such as measurement methodology, power plant age and measurement time of year, in order to better isolate the differences. Analysis must be performed at sub-annual scales in order to isolate consistent measurements at the facilities. These additional attributes combined with a more detailed statistical examination of the differences may uncover the mechanistic drivers of the mismatch between these two datasets, at both the individual and aggregate facility level. Finally, the policy implications of the discrepancies as they relate to the recent EPA rulemaking can inform what steps, if any, must be taken in order to support this and future policymaking on greenhouse gas emissions.

This study asks two questions: (1) what are the differences between these two datasets and what are their statistical properties? (2) Are these differences and the uncertainties they imply large enough to have an impact on policymaking exemplified by the recent EPA final rulemaking on CO_2 emissions from electricity production in the US. Previous attempts to answer the first question have relied on a single year of data and only annual resolution. This, as we will show, does not allow for an accurate determination of which EPA estimates are predominantly continuous emission monitoring (CEMs)-based versus a variety of undocumented substitution methods, critical to a cogent comparison. Furthermore, previous work has assumed Gaussian statistics which is not supported by the distribution of the differences and is therefore, a potentially inaccurate means to assess these differences. Finally, one recent study remains clouded in controversy due to questioned statistical assumptions. No previous studies have attempted to assess how the differences may intersect with recent US policy targeting CO_2 emissions at power production facilities.

In order to answer these questions, we examine the differences between the CAMD and EIA across the 2001–2009 time period in an attempt to better quantify the differences between these datasets. We place these differences within the context of the recent EPA Clean Power Plan (CPP).

2. Methods

The CAMD dataset used in this study is the 'pre-packaged' hourly emissions data for Electrical Generating Units (EGUs) [16]. These data are reported as hourly CO_2 emissions monitored from an emitting stack or through a calculation, based on records of fuel consumption. There are seven categories used to identify the CO_2 emissions estimation method in the CAMD dataset. While we perform some exploration of all the methods [SI Text 2], we focus here on measurements for which a CEMs method is operational on a continuous basis and for those facilities that do not deliver both heat and power. We focus on the CEMs as this is the primary means by which the EPA quantifies power plant emissions.

The EIA dataset which we use here is derived from reporting form 923, which reports monthly data on receipts and cost of fossil fuel, fuel stocks, generation, consumption of fuel for generation, and environmental data at each power plant [17]. We use the supplied CO_2 emission factors to calculate the quantity of CO_2 emitted from the reported consumption data.

In both the CAMD and EIA datasets, every power plant has a unique identifying code, allowing facilities to be organized in matched pairs. Pairs with zero emissions in either or both datasets are removed and the remaining non-zero emitting pairs are used for comparison purposes. Though both datasets include power plants burning other fuel sources, we limit our analysis to fossil fuels only [SI Text 1].

In order to systematically compare the two datasets, we define a series of difference metrics [SI Text 1]. The annual individual difference (ID), is defined as the CO_2 emissions in the CAMD dataset minus the CO_2 emissions in the EIA dataset for each matched power plant. The IRD, is defined as the CO_2 emissions difference at each matched power plant divided by the pair's average value (expressed as a percent). The total difference (TD), between the two datasets is defined as the summed CAMD CO_2 emissions minus the summed EIA CO_2 emissions where the summation occurs over all matched facilities. Similarly, the total relative difference (TRD), is defined as the TD between the two datasets divided by their average value, expressed as a percent.

Similarly, monthly or annual measures of difference can be computed. In the analysis presented here, we use monthly CO_2 emissions, unless specified otherwise.

The Wilcoxon signed-rank test is used to decide whether the median value (e.g., IRD) is significantly different from zero [18]. The Wilcoxon test does not require the data comply with a Gaussian distribution, but it assumes that the data are continuous and symmetrically distributed (no skew) around the median. We chose to use the Wilcoxon test, since our data displays a narrow peak and long tails, which violate the Gaussian distribution assumption.

3. Results

3.1. Differences between the two datasets

Figure 1 shows standardized frequency distributions of the IRD values of monthly CO_2 emissions between matched power plants for the year 2001 and 2009 in addition to standard normal distributions. The IRD distributions exhibit a narrower central peak and longer distribution tails compared to standard normal distributions. Furthermore, the IRD distributions exhibits asymmetrical qualities, highlighted by the differences between the mean and median values. Therefore, we avoid the use of standard deviation in describing the distribution and instead rely on direct quartile and quintile metrics.

Figure 2 shows the TRD, the median IRD values and the three inner quintiles for each of the 9 years in the 2001–2009 time period. The median IRD represents the median of large positive and negative monthly differences and this can be seen from the boundaries of the quintiles shown in figure 2. For example, in the year 2001 the median IRD value is +2.1%, yet the upper (lower) fifths of the distribution show IRD values that are greater (less) than +10.8% (−6.2%). The median IRD value in all years is positive indicating that the majority of matched facilities have larger values in the CAMD versus the EIA data. The median IRD values initially increase, reaching a value of +3.2% then decline to values less than +1% from 2006 and onward. The IRD distributions also tend to be wider in the earlier portion of the record suggesting a larger number of facilities with large IRD values. For example, by 2009, the upper (lower) fifths of the distribution show IRD values that are greater (less) than +5.9% (−6.1%).

The TD between the matched facilities also is largest and increasing in the first 3 years of the time series up to a maximum value of 2.2% in 2003, after which it declines to values less than ±0.6%. These results demonstrate that although the mean differences are not large (rarely exceeding 3%), this is the result of large negative and positive IDs that cancel in the aggregate.

3.2. Mapping maximum differences

The differences at each of the individual matched facilities can be examined in space. Figure 3 presents both the maximum ID and maximum IRD at each

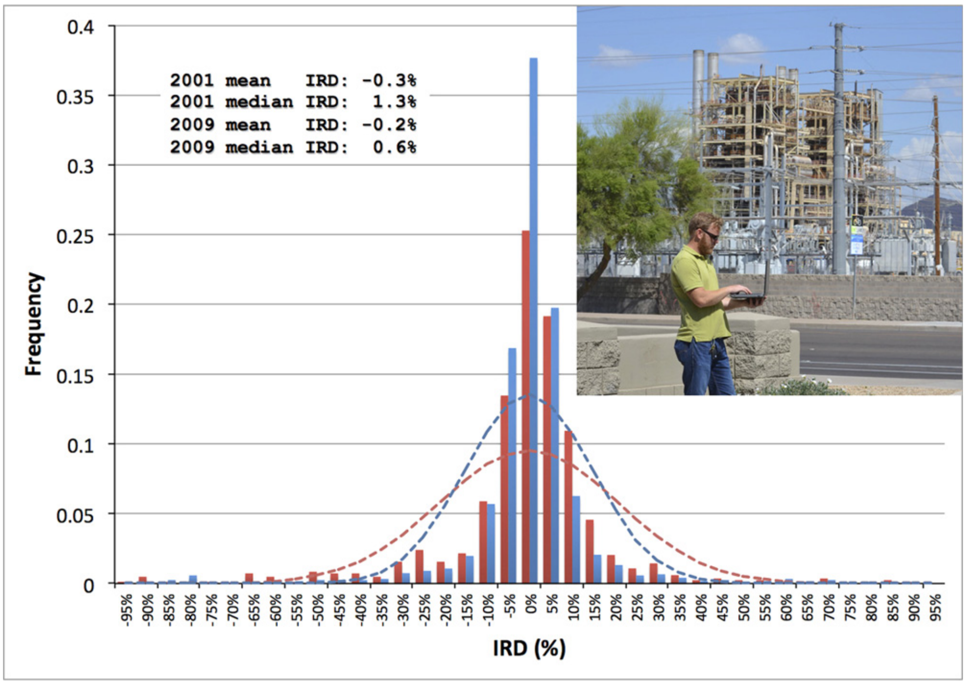

Figure 1. Standardized frequency distribution of the annual CO_2 emissions individual relative difference (IRD) values of matched power plants for the years 2001 (red) and 2009 (blue). Standard normal distributions for each are also shown (dashed lines) in addition to mean and median values.

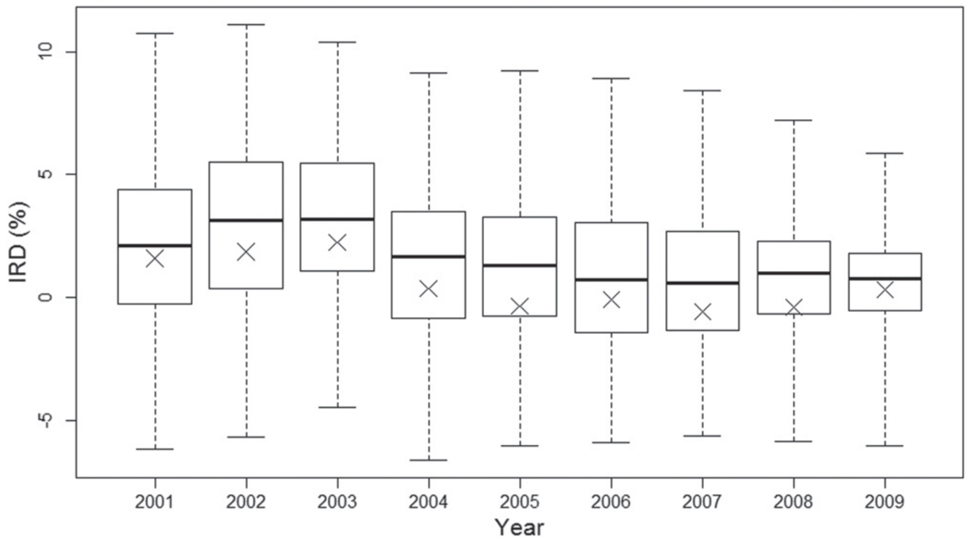

Figure 2. Median of the CO_2 emissions individual relative difference (IRD—central line symbol) values, CO_2 emissions total relative difference (TRD—X symbol) values, and quintile distribution boundaries of the CO_2 emissions IRD values at matched power plants for each year in the 2001–2009 time period. Box represents central quintile of cumulative IRD distribution. Whiskers denote the next quintile. Facility pairs include only power plants utilizing CEMs continuously over a month and do not deliver both heat and power. All median IRD values are statistically significant at the $p < 0.01$ level using a Wilcoxon test (determines whether median IRD is significantly different from zero).

facility across all 9 years of monthly differences. The median of the absolute ID values across the facilities in the eastern half of the US (using -95 longitude as dividing line) is nearly 40% larger than the median of the absolute ID values in the western half. The IRD values, by contrast, show somewhat less geographic dependence with the median of the absolute IRD values nearly 20% larger in the East than the West.

The facilities that occupy the top differences shown in figure 3(a) are candidates for deeper onsite evaluation. We define these top facilities as those which emerge repeatedly within the top ID values in each of the 9 years. This reflects those facilities with both large ID values and a persistent presence among the large ID values in each of the 9 years of matched data. Table 1 lists these facilities and the number of

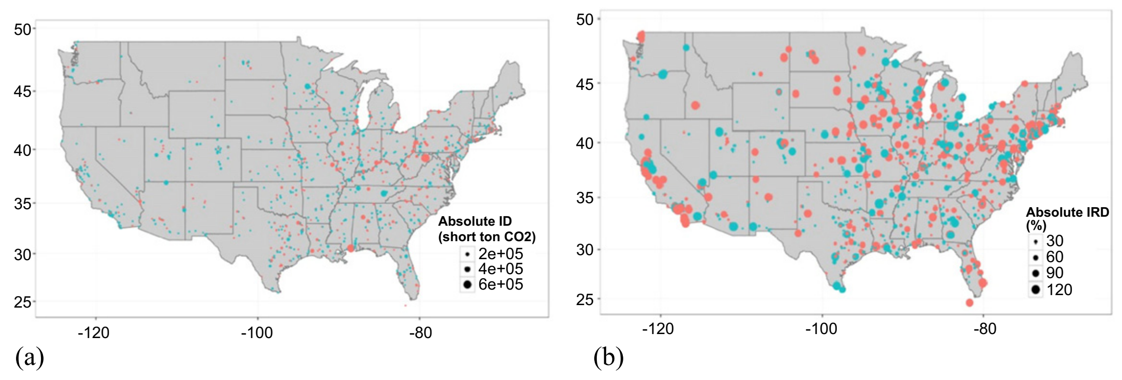

Figure 3. Maximum difference values at each of the matched facilities. (a) Maximum ID value; (b) maximum IRD value. Red symbols: >0, green symbols: <0.

Table 1. Individual facilities with consistently large ID values.

Facility name	ORISPL code	State	Frequency
John E Amos	3935	WV	4
Cumberland	3399	TN	3
Bruce Mansfield	6094	PA	5
Intermountain	6481	UT	5
Mount Storm Power Station	3954	WV	4
Rockport	6166	IN	5
Gen J M Gavin	8102	OH	3
Pleasant Prairie	6170	WI	3
Clifty Creek	983	IN	3
Labadie	2103	MO	6
Navajo Generating State	4941	AZ	3
Powerton	879	IL	5

times they ranked in the top 12 differences in the 9 years. A complete list of the top ten in each of the nine years is provided in the supplementary information [see SI Text 3].

4. Discussion

Differences between these two United States power plant energy/emissions datasets have implications for policy and decisionmaking though the relevance is dependent upon the spatial scale and scope of policy purpose. For example, the United States greenhouse gas inventory, used domestically and internationally, uses power plant data derived from EIA fuel statistics [19]. In the national aggregate, these data are consistent with the EPA power plant CO_2 emissions estimation (1% to 2%). However, this small difference masks an important element when comparing the underlying datasets. First, though the aggregate values are nearly identical between the two datasets, this is the result of large positive and negative differences, which cancel in the aggregate. For example, on those power plants for which an EPA direct emission monitor is actively sampling, differences between these and the EIA estimates based on the consumed fuel, are greater than −6% for 20% of the facilities and greater than +9% for 20%. This increases to −13% and +14% for the outer 20% of the facilities.

Differences this large at the individual facility level raise concerns regarding policy operationalized at the US subnational level for which examples already exist at state and municipal levels [20]. At these scales, the choice of dataset will have significant implications on baseline emissions and policy outcomes. More importantly, actual emissions may indeed be accurately portrayed in one or the other of these datasets and hence, the wrong choice, could lead to biased outcomes and misguided policy.

Recently, the United States Environmental Protection Agency announced the implementation of the Clean Power Plan (CPP) for existing power plants [12, 21]. The rule establishes state-specific targets for lowering the average emission rate (lbs CO_2/MWhr) from a state's EGUs. The targets are to be achieved by the year 2030 and represent reductions relative to the year 2012. The rule establishes 'interim' goals that states must meet over the 2020–2029 time period to ensure compliance with the final target in 2030. For example, the state of Arizona must lower its state-average emission rate from a 2012 mean value of 1551 lbs CO_2/MWhr to a final level of 1031 lbs CO_2/MWhr by the year 2030 with an interim target of 1173 lbs CO_2/MWhr (measured as an average of the 2020–2029 time period). In percentage terms, the 2030 target is equivalent to a 34% reduction.

In order to demonstrate compliance with the targets established in the rule, state's will have to measure or otherwise estimate their state-average EGU CO_2 emission rate. Such measurements or estimation procedures will depend upon the mix of policies and measures adopted to meet their target. For example, demand side energy efficiency improvements will require a means to estimate the amount of electricity demand obviated. Expansion of nuclear or renewable electricity supply will require estimation of the amount of zero-carbon electrical generation. However, because fossil fuel EGUs will remain a

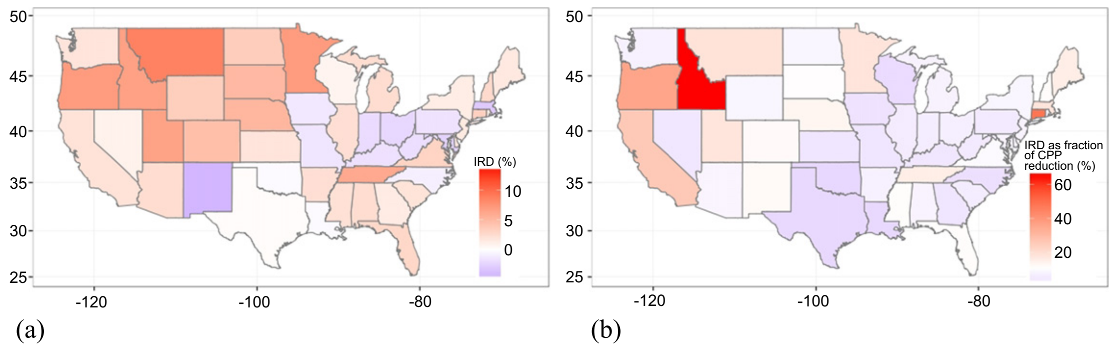

Figure 4. State-level CO_2 emission differences in the context of the EPA CPP. (a) IRD values for the state-level aggregation of the matched facilities in the CAMD versus EIA data (red colors >0%, blue colors <0%); (b) these IRD values as a percent of the EPA power plant rule reduction amounts (red colors >10%, blue colors <10%).

component of all state's energy supply and is the part of the US emitting landscape targeted by the EPA rule, it will be essential to measure or otherwise estimate the amount of CO_2 emitted at the power plant level.

We have aggregated the individual facility emissions in the two datasets by state and calculated the state-level IRD values (figure 4(a)). As with the national facility-level results, the state-level median IRD is +1.9% with 20% of the values greater/less than −1.2/+4.2%. Though these state-level differences are smaller than the individual facility differences owing to in-state cancelation of positive and negative differences, the state-level differences must be cast in the context of the proposed EPA Plan reduction amounts. Hence, in figure 4(b), we show the magnitude of the state-level IRD values as a fraction of the percentage reductions associated with the EPA rule. For example, the state of Minnesota must meet a state-wide carbon emission rate of 1213 lbs CO_2/MWhr in its fossil-based power plants by the year 2030. This constitutes a 39.7% reduction from its 2012 state-wide emission rate of 2013 lbs CO_2/MWhr. With an estimated IRD for the state of 7.1%, this suggests that the state of Minnesota could have an actual emission rate 18% (7.1/39.7) higher/lower than the emission rate they assume they have when attempting to meet their target amount. Worse yet, over one-third of the states show potential over/under-estimated percentage amount in the double-digits.

The inability of state's to reliably use the existing monitoring systems for power plant emissions suggests the need for further investigation into more reliable monitoring, an assessment of which of these datasets is more accurate, or both. Such an assessment is critical for planning, implementing and verifying the emission rate goals in the EPA rule. Given the potential uncertainties identified in this study, controversy may arise on questions of compliance with the EPA regulations.

5. Conclusions

The carbon dioxide emission from US power plants represents roughly 40% of the national fossil fuel CO_2 emissions budget. Hence, these point sources are a critical element in policies aimed at lowering greenhouse gas emissions. Two datasets provided by two different US agencies have tracked the CO_2 emissions from US power plants: the Department of Energy's EIA and the Environmental Protection Agency. Analysis of the matched plants in these two datasets shows large offsetting differences in estimated CO_2 emissions which are masked when aggregated to a national sum. For example, over the 2001–2009 time period, nearly one-half of the facilities have monthly emission differences that exceed roughly ±6% and one-fifth exceed roughly ±13%. When aggregated to the state-level, differences remain large relative to the recently announced US CPP which aims to lower the CO_2 emissions at fossil fuel-based power plants in the US by the year 2030. When these differences are cast as a proportion of the percent reduction, one-third of US states have percent proportions greater than 10%, differences this large at the individual facility and state-level raise critical concerns about the ability of states to comply with the rule or the reality of the actual emission reductions.

Acknowledgements

This research was supported by the National Science Foundation on CAREER award 0846358. We thank Jeffrey Quick of the Utah Geological Service for helpful comments. We thank EPA staff for additional input on CEMs methods and documentation.

References

[1] IPCC 2013 *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed T F Stocker *et al* (Cambridge: Cambridge University Press) p 1535

[2] Gurney K R *et al* 2007 Research needs for process-driven, finely resolved fossil fuel carbon dioxide emissions *EOS Trans. Am. Geophys. Union* **88** 542–3
[3] Gurney K R, Mendoza D, Zhou Y, Fischer M, de la Rue du Can S, Geethakumar S and Miller C 2009 The vulcan project: high resolution fossil fuel combustion CO_2 emissions fluxes for the United States *Environ. Sci. Technol.* **43** 5335–541
[4] Gurney K R, Razlivanov I, Song Y, Zhou Y, Benes B and Abdul-Massih M 2012 Quantification of fossil fuel CO_2 at the building/street scale for a large US city *Environ. Sci. Technol.* **46** 12194–202
[5] Miller J B *et al* 2012 Linking emissions of fossil fuel CO_2 and other anthropogenic trace gases using atmospheric $^{14}CO_2$ *J. Geophys. Res.* **117** D08302
[6] Schneising O, Heymann J, Buchwitz M, Reuter M, Bovensmann H and Burrow J P 2013 Anthropogenic carbon dioxide source areas observed from space: assessment of regional enhancements and trends *Atmos. Chem. Phys.* **13** 2445–54
[7] Pacala S *et al* 2010 *Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements* National Research Council
[8] Duren R M and Miller C E 2011 Towards robust global greenhouse gas monitoring *Greenhouse Gas Meas. Manage.* **1** 80–4
[9] D'Avignon A, Carloni F A, La Rovere E L and Dubeux C B S 2010 Emission inventory: an urban public policy instrument and benchmark *Energy Policy* **38** 4838–47
[10] Hoesly R, Blackhurst M, Matthews H S, Miller J F, Maples A, Pettit M, Izard C and Fischbeck P 2012 Historical carbon footprinting and implications for sustainability planning: a case study of the pittsburgh region *Environ. Sci. Technol.* **46** 4283–90
[11] Petron G, Tans P, Frost G, Chao D and Trainer M 2008 High resolution emissions of CO_2 from power generation in the USA *J. Geophys. Res.* **113** G04008
[12] Federal Register 2015 *40 CFR Part 60, Carbon Pollution Emission Guidelines for Existing Stationary Sources: Electric Utility Generating Units; Final Rule* Environmental Protection Agency
[13] Ackerman K V and Sundquist E T 2008 Comparison of two US power-plant carbon dioxide emissions data sets *Environ. Sci. Technol.* **42** 5688–93
[14] Quick J 2014 Carbon dioxide emission tallies for 210 US coal-fired power plants: a comparison of two accounting methods *J. Air Waste Manage. Assoc.* **64** 73–9
[15] Gurney K R, Huang J and Coltin K 2014 Comment on quick, J C (2014) carbon dioxide emission tallies for 210 US coal-fired power plants: a comparison of two accounting methods J. Air Waste Manage. Assoc. 64: 73–79 *J. Air Waste Manage. Assoc.* **64** 1215–7
[16] Air Markets Program Data 2012 Pre-packaged data (http://ampd.epa.gov/ampd/) (accessed 28 May 2012)
[17] Department of Energy/Energy Information Administration 2003 *Electric Power Monthly March 2003* Energy Information Administration, Office of Coal, Nuclear, and Alternate Fuels, US Department of Energy, Washington D.C. 20585
[18] Gibbons J D and Chakraborti S 2011 *Nonparametric Statistical Inference* (Berlin: Springer)
[19] USEPA 2012 *Inventory of US Greenhouse Gas Emissions and Sinks: 1990–2010* US Environmental Protection Agency, 1200 Pennsylvania Ave., NW, Washington, DC 20460, USA, 15 April
[20] Lutsey N and Sperling D 2008 America's bottom-up climate change mitigation policy *Energy Policy* **36** 673–85
[21] Gurney K R 2015 What is the role for carbon cycle science in the proposed EPA power plant rule? *Earth Perspect.* **2** 1–6