4th International Workshop
on Uncertainty in Atmospheric Emissions
7-9 October 2015, Krakow, Poland

PROCEEDINGS

4th International Workshop on Uncertainty in Atmospheric Emissions

7–9 October 2015, Kraków, Poland

PROCEEDINGS

Warszawa 2015

4th International Workshop on Uncertainty in Atmospheric Emissions
7- 9 October 2015, Cracow, Poland

Printed from the material submitted by the authors.

ISBN 83-894-7557-X
EAN 9788389475572

About the Workshop

The assessment of greenhouse gases and air pollutants (indirect GHGs) emitted to and removed from the atmosphere is high on the political and scientific agendas. Building on the UN climate process, the international community strives to address the long-term challenge of climate change collectively and comprehensively, and to take concrete and timely action that proves sustainable and robust in the future. Under the umbrella of the UN Framework Convention on Climate Change, mainly developed country parties to the Convention have, since the mid-1990s, published annual or periodic inventories of emissions and removals, and continued to do so after the Kyoto Protocol to the Convention ceased in 2012. Policymakers use these inventories to develop strategies and policies for emission reductions and to track the progress of those strategies and policies. Where formal commitments to limit emissions exist, regulatory agencies and corporations rely on emission inventories to establish compliance records.

However, as increasing international concern and cooperation aim at policy-oriented solutions to the climate change problem, a number of issues circulating around uncertainty have come to the fore, which were undervalued or left unmentioned at the time of the Kyoto Protocol but require adequate recognition under a workable and legislated successor agreement. Accounting and verification of emissions in space and time, compliance with emission reduction commitments, risk of exceeding future temperature targets, evaluating effects of mitigation versus adaptation versus intensity of induced impacts at home and elsewhere, and accounting of traded emission permits are to name but a few.

The *4th International Workshop on Uncertainty in Atmospheric Emissions* is jointly organized by the *Systems Research Institute of the Polish Academy of Sciences*, the Austrian-based *International Institute for Applied Systems Analysis*, and the *Lviv Polytechnic National University*. The 4th Uncertainty Workshop follows up and expands on the scope of the earlier Uncertainty Workshops – the *1st Workshop* in 2004 in Warsaw, Poland; the *2nd Workshop* in 2007 in Laxenburg, Austria; and the *3rd Workshop* in 2010 in Lviv, Ukraine.

Steering Committee

Rostyslav BUN (Lviv Polytechnic National University, UA)
Matthias JONAS (International Institute for Applied Systems Analysis, AT)
Zbigniew NAHORSKI (Polish Academy of Sciences, PL) – Chair

Scientific Committee

Evgueni GORDOV (Siberian Center for Environmental Research & Training, RU)
Piotr HOLNICKI-SZULC (Polish Academy of Sciences, PL)
Joanna HORABIK-PYZEL (Polish Academy of Sciences, PL)
Olgierd HRYNIEWICZ (Polish Academy of Sciences, PL)
Katarzyna JUDA-REZLER (Warsaw University of Technology, PL)
Petro LAKYDA (National University of Life and Environmental Sciences of Ukraine, UA)
Myroslava LESIV (Lviv Polytechnic National University, UA)
Gregg MARLAND (Appalachian State University, USA)
Sten NILSSON (Forest Sector Insights AB, SE)
Tom ODA (Univ. Space Research Association, NASA Goddard Space Flight Center, USA)
Stefan PICKL (Universität der Bundeswehr München, Germany)
Elena ROVENSKAYA (International Institute for Applied Systems Analysis, AT)
Kazimierz RÓŻAŃSKI (AGH University of Science and Technology in Cracow, PL)
Dmitry SCHEPASCHENKO (International Institute for Applied Systems Analysis, AT)
Anatoly SHVIDENKO (International Institute for Applied Systems Analysis, AT)
Jacek SKOŚKIEWICZ (National Centre for Emissions Management, PL)
Philippe THUNIS (EC Joint Research Centre Ispra, EU)
Marialuisa VOLTA (University of Brescia, IT)

Local Organizing Committee

Joanna HORABIK-PYZEL
Jolanta JARNICKA - Chair
Weronika RADZISZEWSKA
Jörg VERSTRAETE

Uncertainty in gridded CO_2 emissions estimates

Susannah Hogue[1], Eric Marland[2], Robert J. Andres[3], Gregg Marland[4], Charles Robison[5], Dawn Woodard[6]

[1] Department of Mathematical Sciences, Appalachian State University, Boone NC, USA. hogues@appstate.edu
[2] Department of Mathematical Sciences and Research Institute for Environment, Energy, and Economics, Appalachian State University, Boone NC, USA. marlandes@appstate.edu
[3] Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge TN, USA. andresrj@ornl.gov
[4] Research Institute for Environment, Energy, and Economics, Appalachian State University, Boone NC, USA. marlandg@appstate.edu
[5] Department of Geography and Planning, Appalachian State University, Boone NC, USA. robisonch@appstate.edu
[6] Department of Earth Systems Science, University of California, Irvine CA, USA. dawnwoodard@gmail.com

Abstract

We are interested in the spatial distribution of fossil-fuel-related emissions of CO_2, but it is important to understand the uncertainty in emissions estimates. Uncertainty is introduced in the magnitude and location of large point sources, the magnitude and distribution of non-point sources, and from the use of proxy data to characterize emissions. For the U.S. we develop estimates of the contribution of each component. At 1 degree resolution, in most grid cells, the largest contribution to uncertainty comes from how well the distribution of the proxy (population density) represents the distribution of emissions. In other grid cells the magnitude and location of large point sources make the major contribution. Uncertainty is strongly scale-dependent with uncertainty increasing as grid size decreases. Uncertainty for one degree grid cells is typically on the order of +/- 150% but this is perhaps modest in a data set where emissions per grid cell vary over 8 orders of magnitude.

Keywords: U.S. CO_2 emissions, gridded emissions, large point sources, proxy data

1. Introduction

There is a wide range of interest (both geochemical and geopolitical) in geographically explicit inventories of the sources and sinks of the greenhouse gas CO_2. It is a challenge to estimate sources and sinks in a spatially-explicit context and to best characterize the location and magnitude of emissions and sinks we would like to estimate also the associated uncertainty. Current gridded inventories of emissions from fossil-fuel use and industrial processes rely heavily on related, proxy and re-purposed data. In the following analyses we refine and combine the components of uncertainty and discuss them in the context of the widely-used Carbon Dioxide Information Analysis Center [1] gridded inventory for fossil-fuel related emissions from the U.S. (see also [2]).

Few studies have explored the uncertainty of global-scale, grid-level emissions datasets. Rayner et al. [3] noted that "none of the pointwise fossil emission products available today include" estimates of uncertainty and then estimated that for their dataset "uncertainties can be as high as 50% at the pixel level". They also pointed out, importantly, that uncertainties for nearby pixels are not independent because, for example, the uncertainty for any given grid space includes consideration that a large point source might be only slightly displaced and the total for the ensemble of cells is constrained by national data. Rayner et al. emphasize that "using the uncertainty of this pointwise map alone in an inversion is a serious error since it assumes independence of errors."

2. Methods and analysis

Large point sources make up a large percentage of anthropogenic carbon dioxide emissions for the U.S. and for other industrialized countries [4]. In 2010 one third of U.S. emissions were reported from only 311 sites of large point sources [5]. As soon as the first few latitude and longitude data points from these data sets were typed into Google Earth many of these point sources were not observed at their reported locations. We have to deal with both magnitude and locational uncertainty. Total uncertainty in emissions from any geographic grid space thus has to reflect uncertainty in small or areal sources and in both the magnitude and location of large point sources.

Woodard et al. [6] have developed one key component of what we need to quantify the spatially explicit uncertainty in gridded inventories of CO_2 emissions – an approach for dealing with the uncertainty in the locations of large point sources. Also many gridded inventories exist that document ground level sources of anthropogenic emissions of CO_2 for the U.S. and the globe and these inventories use a variety of top-down and bottom-up methods to geographically distribute emissions that are not attributed to large point sources. Each of the top-down inventories uses some sort of proxy, such as population density or satellite-observed nightlights, to help distribute emissions totals from a large (national or state) scale down to the level of grids as small as 0.1 degrees on a side. Some of the inventories use multiple proxies to take advantage of their differing characteristics. Using proxy data, while necessary, can result in the misallocation of emissions values both spatially and temporally. Hutchins et al. [7] show that the differences among existing data sets increase as grid size is decreased. To address these issues, we have taken the first steps toward calculating the total uncertainty for a CDIAC-like inventory for the U.S. at the 1-degree grid scale. Data here are estimates of annual emissions for the year 2009.

In the gridded CDIAC inventory [1] data on population density are used as a proxy for the spatial distribution of all emissions within a country. For this analysis, we have removed emissions from electric power plants from the country total prior to using the population proxy to distribute the remaining national emissions. The power plants, with magnitudes and locations from EPA's eGRID dataset [5], were then added back to give total emissions in each grid cell. The emissions inventory discussed here is thus comprised of two components, power plant emissions from the eGRID dataset and all remaining national emissions distributed using population density as a proxy. These remaining emissions do contain some additional, large industrial sources of CO_2, but reporting to the EPA GHG Reporting Program [8] shows that in 2010, 73% of emissions from sources greater than 25,000 metric tons of CO_2 equivalent were from power plants. Comparable data on industrial sources are not available outside of the U.S. and for the purposes of this study we assume these industrial sources to be part of the areal sources of emissions (hereafter "non-point sources").

There are thus 6 components of uncertainty that need to be combined for an estimate of total uncertainty for the cells in the modified CDIAC database:

- Uncertainty in total national emissions
- Magnitude uncertainty for large point sources
- Spatial uncertainty for large point sources
- Magnitude uncertainty of the population proxy

- Spatial uncertainty for the population proxy
- Uncertainty in using population density as a proxy for emissions

For this analysis all calculations are based on one-sigma uncertainty. In combining the component uncertainties we assume that, except for the uncertainty in the national total, each of the components are independent of each other and they are therefore combined in Euclidean fashion (the square root of the sum of the squares) on a grid cell by grid cell basis. The uncertainty in the national total is passed to all components equally across grid cells. In this analysis we use data for emissions in 2009 as published by CDIAC in 2013. Temporal uncertainty would need to be considered in developing a time series of emissions.

Uncertainty in total national U.S. emissions is estimated at 2.5% (one sigma). This value is based on comparisons with other inventories, U.S. EPA analyses, and literature research of the components which combine to calculate total national U.S. emissions. (see also Andres et al. [2]). According to the U.S. EPA [9] the 95% confidence interval for total U.S. CO_2 emissions from fossil fuel combustion is -2% to +5%. Instead of using this asymmetric value we take the symmetric value of ±5%, and since this is two standard deviations about the mean and our computations are all based on one standard deviation, the estimated national error used in our computations is ±2.5%.

A random sample of 500 large point sources from the US EPA eGRID data set was taken in order to find the exact locations of the power plant discharges. We used Google Earth satellite imagery to identify the point sources and to verify each latitude and longitude. With spatial information from Google Earth, the distance between the actual location and the reported location was computed for each point source in the sample. The maximum distance from the reported location to the observed location of a point source was approximately 106 km. The mean distance from the reported location for all of the sample point sources (excluding zero) was 1.97 km. The mean distance from the reported locations for all of the point sources in the sample was 0.84 km. The latter value was then used as the mean spatial uncertainty. The spatial uncertainty for the top 81 emitters was larger than for the random sample of points. It was found that 60% were farther than 1km from the reported location. The mean difference in location was 7.94 km and the maximum spatial difference was about 122 km.

The information gathered from the 500-item random sample suggested that the differences between discharge locations and eGRID reported locations might be attributed to: 1.) differences between the plant site and the exact location of the CO_2 discharge stack, 2.) use of default locations in the EPA database, such as the centroid of a county, when the initial report to EPA did not include plant coordinates, 3.) typographical errors, 4.) reporting the location of a company office or mailing address instead of the plant site, and 5.) dealing with the existence of multiple stacks on the same site.

The locations of power plants are not part of a continuous distribution and therefore most traditional statistical methods do not work well in dealing with the uncertainty in their emissions. The discrete, or binary, nature of the locations (a plant either is in a given grid space or it is not) spurred the creation of a new method for dealing with the likely locations and the uncertainty in the emissions from power plants and the development of a new statistic, PSUM = Point Source Uncertainty Measure, [6] which we treat as a standard deviation in the analyses here. The uncertainty results are scale dependent, as with any spatial uncertainty. For a given locational uncertainty, the larger the grid cell the greater the probability that the point source will actually be found in the reported cell. Monte Carlo analyses were run using the magnitude of emissions and the reported location for each point source as well as the calculated mean spatial uncertainty and the size of the geographic grid cells. A resulting grid of simulated means effectively distributes the reported CO_2 emissions from a point source to surrounding cells based on the fraction of the total number of simulation executions that fell in each cell.

The magnitude uncertainty for emissions from large point sources is taken to be a constant ±10.62% (one sigma) [10]. This number was derived by comparing data collected on smokestack emissions of U.S. electric power plants with emissions calculated from fuel deliveries at the same plants. The value 10.62 was the mean of the difference of the two measurements.

LandScan is a recently developed global data set [11] that estimates the average locations where people actually are rather than where their home location is. Landscan was first produced in 1998 and data sets for 2000-2012 are now available. CDIAC has contemplated use of LandScan population data but has not yet made the conversion. The CDIAC gridded CO_2 emissions data set relies, however, on a 1984 population distribution data base from the Goddard Institute for Space Studies [12]. While this GISS data set allows an estimation of the spatial distribution of CO_2 emissions over a long time series, an additional contribution to uncertainty results from the changes in urbanization and population distribution that have occurred since 1984. Although the GISS data are used as the proxy for distributing emissions, we use LandScan characteristics here to illustrate the uncertainty that could be achieved for the post-2000 time period.

Magnitude uncertainty in LandScan for the U.S. was assumed to be comparable to the estimates of uncertainty derived by the U.S. Census Bureau at the same spatial scale [13]. LandScan does not currently have any published estimates for uncertainty but we assume that it is very low in the U.S. As in all of the data sets used here, the uncertainty will vary by country or region in a global analysis. The uncertainty estimate provided by the U.S. Census Bureau is 0.01%. Spatial uncertainty in LandScan was estimated by looking at the changes incurred as a result of small shifts in the cell boundaries. We took the LandScan data set and distributed CO_2 emissions proportional to the population density values associated with each grid cell. We then shifted the grid cells by 10 kilometers (approximately one tenth of a grid cell in the central U.S.) in each direction (N, S, E, and W) so that each grid cell contained successively one tenth of each of the four surrounding cells. This effectively creates a weighted sum in which the central cell emissions value is weighted by 90% and the cell that is shifted towards the center is weighted by the remaining 10%. A weighted sum was computed for each of the four shifts that occurred. The standard deviation for the resulting four weighted sums was then computed and stored as the uncertainty value within the central cell.

In order to characterize the uncertainty associated with using population density as a proxy for CO_2 emissions we started with the per capita emissions in each state [14] (with data on large point sources removed) and the mean number of grid spaces per state. We calculated the standard deviation in per capita emissions from non-point sources at the state level and took this as a measure of the variability in the relationship between population density and emissions density at that scale. We assumed that the variability among states provides a measure of the variability at the grid level within states. This provides enough information that we can back-calculate to estimate the standard deviation in emissions by grid cell attributed to the population proxy. Then we can use this standard deviation as the uncertainty estimate for using population density as a proxy for emissions at the grid cell level.

3. Discussion and conclusions

The combination of the six aspects of uncertainty produces a total uncertainty, by gridspace, for a hypothetical, modified CDIAC dataset. Figure 1 shows these results as a percent of total emissions in each grid space. Recall that these numbers apply to a hypothetical data set – one in which 1.) the data on large point sources have been substituted for an equal quantity of emissions that were previously distributed according to population density and 2.) the data on population density have an uncertainty attributed to data from the U.S. Census Bureau at the same scale. In both cases we expect that the best achievable uncertainty will be much higher in many other countries where the data on population and large point sources have greater uncertainty.

The scale shown on Figure 1 is quite high. This was not unexpected because of how much uncertainty there is for the exact location and magnitude for CO_2 sources at this scale. Figure 1 shows that the uncertainty associated with the modified CDIAC data set is consistently around 150% of the emissions total for each grid space. Recall that these estimates of uncertainty are for a modified CDIAC data set where we have now isolated large point sources before using population density to distribute the remaining emissions from non-point sources. We have also treated the population density data as though they had been derived from Landscan, thus avoiding the shifts in population density that have occurred since construction of the Goddard Institute for Space Studies data set for 1984. Recall too that this uncertainty is for individual grid cells of 1 degree scale and is very scale dependent. There is strong correlation among grid spaces because of the spatial uncertainty about the exact placement of large point sources and because the national total is a defined constant.

Total uncertainty is the combination of all of the components, but we also learn something of the role that each component takes in forming the whole. With an understanding of the relative magnitude of each of the pieces, and the locational characteristics of where each component is large, we can try to target specific efforts to best reduce the total uncertainty. Table 1 provides summary statistics on the different components of uncertainty. Table 1 indicates that proxy uncertainty has the highest mean percentage of all the components. In particular a full 52% of grid cells have 90% of their uncertainty coming from proxy uncertainty. The implication here is that in the majority of grid cells, reduction of uncertainty can only be done by addressing uncertainty in the proxy relationship. This means that we must obtain a better understanding of the relationship between the proxies we use and the emissions they are meant to represent. Contributions from large point sources often dominate uncertainty for the grid cells where large point sources are present. And, the values here depend very much on the geographic scale. Uncertainty will increase for many reasons if the grid size is decreased without reducing the parameters of spatial uncertainty.

Our efforts to systematically estimate the uncertainty in a gridded data set of CO_2 emissions suggest that the uncertainty is quite high in the U.S. and it is probably higher in many countries where data on large point sources and the distribution of population are less well documented. Uncertainty will increase as the geographic scale is decreased. While data users need to appreciate the data uncertainty, the best data are probably suitable for many purposes. The analyses suggest that at 1-degree latitude/longitude resolution the current uncertainty (one standard deviation) by grid space in the U.S. is on the order of +/- 150%. Taking this analysis to a global scale will require additional analysis to characterize spatial uncertainty for each country or group of similar countries.

Acknowlegement:

Portions of this work were supported by NASA CMS grant NNH11ZDA001N-CMS.

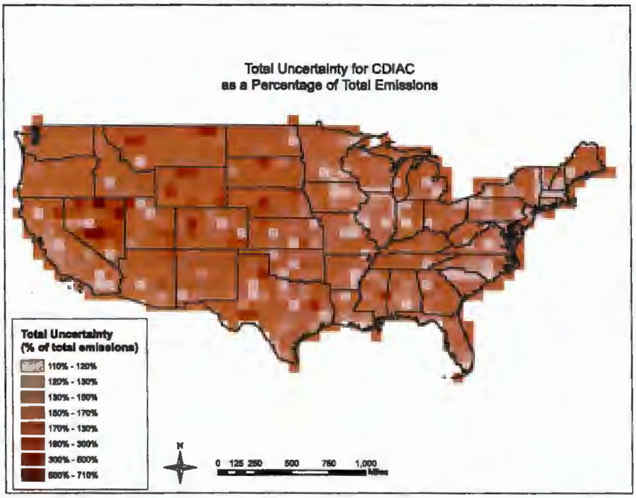

Figure 1. Uncertainty by grid space shown as a percentage of total emissions at 1-degree resolution. Areas shown with very high uncertainty are often a result of cities with abrupt changes in population density. Excluding these few areas of very high uncertainty we can see that the overall uncertainty in grid spaces is on the order of 1.5 times the total emissions. This is for a hypothetical, modified version of the CDIAC data set for 2009 (see text).

Table 1. By grid cell, a breakdown of each component of uncertainty with its summary statistics. Uncertainty in the national total is not included since it affects each of the grid cells equally. Note that the country borders create problems in that emissions may or may not occur even if a grid cell is designated as predominantly ocean, and some of the zero values lie along the eastern shoreline of the U.S. This is one of the challenges of cropping a global data set to a single country for a targeted analysis. Values are given as the percent uncertainty in a single grid cell.

	Min	1st Quart.	Median	Mean	3rd Quart.	Max
Magnitude Uncertainty, large point sources	0.00	0.00	13.96	41.70	94.88	110.60
Spatial Uncertainty, large point sources	0.00	0.00	12.62	37.69	85.77	100.00
Magnitude Uncertainty, LandScan	0.00	0.00	0.01	0.01	0.01	0.01
Spatial Uncertainty, LandScan	0.00	0.09	0.55	10.34	2.93	673.00
Proxy Uncertainty	0.00	22.99	141.20	100.70	161.50	161.50
Total Uncertainty, Emissions Data at Grid Cell Level	112.9	138.3	153.6	154.2	166.4	712.9

(Quart.=Quartile)

References Cited

[1] Andres, R.J., T.A. Boden, and G. Marland, 2013. Annual fossil-fuel CO_2 emissions: mass of emissions gridded by one degree latitude by one degree longitude. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tenn., U.S.A. doi 10.3334/CDIAC/ffe.ndp058.2013

[2] Andres R.J., T.A. Boden, and D. Higdon, 2014. A new evaluation of the uncertainty associated with CDIAC estimates of fossil fuel carbon dioxide emission. Tellus B, 66, 23616. doi:10.3402/tellusb.v66.23616.)

[3] Rayner, P.J., M.R. Raupach, M. Paget, P. Peylin, and E. Koffi, 2010. A new global gridded data set of CO_2 emissions from fossil fuel combustion: methodologies and evaluation, J. of Geophysical Research v. 115, pp. D19306, doi: 10.1029/2009JD013439.

[4] Singer, A., M. Branham, M. Hutchins, J. Welker, D. Woodard, C. Badurek, E. Marland, and G. Marland, 2014. ``The role of CO_2 emissions from large point sources in emissions totals, responsibility, and policy". Environmental Science and Policy, 44:190-200.

[5] U.S. EPA, 2014. Clean Energy: eGRID, ninth edition with 2010 data. U.S. Environmental Protection Agency, http://www.epa.gov/cleanenergy/energy-resources/egrid/, released February 24, 2014.

[6] Woodard, D., M. Branham, G. Buckingham, S. Hogue, M. Hutchins, R. Gosky, G. Marland, and E. Marland, 2015. A spatial uncertainty metric for anthropogenic CO_2 emissions. Greenhouse Gas Measurement and Management, 4:1-22. doi:10.1080/20430779.2014.1000793

[7] Hutchins, M.G.H., J.D.C. Colby, G. Marland, and E. Marland, 2015. A comparison of five high-resolution spatially-explicit fossil fuel carbon dioxide emissions inventories, in review.

[8] U.S. EPA, 2013. GHGRP 2011: Reported Data, U.S. Environmental Protection Agency, Greenhouse Gas Reporting Program. http://www.epa.gov/ghgreporting/

[9] U.S. EPA, 2015. U.S. Greenhouse Gas Inventory Report: 1990-2013, U.S. Environmental Protection Agency, http://epa.gov/climatechange/ghgemissions/usinventoryreport.html

[10] Quick, J.C., 2014. Carbon dioxide emission tallies for 210 U.S. coal-fired power plants: a comparison of two accounting methods. J. Air and Waste Management Assoc. 64:73-79.

[11] Oak Ridge National Laboratory, 2014. LandScan, 2014. http://web.ornl.gov/sci/landscan/, retrieved June 10, 2015.

[12] Andres, R.J., G. Marland, I. Fung, and E. Matthews, 1996. A 1° x 1° distribution of carbon dioxide emissions from fossil fuel combustion and cement manufacture, 1950-1990. Global Biogeochemical Cycles10(3): 419-429.

[13] U.S. Census Bureau, 2011. Population distribution and change: 2000 to 2010: Census Briefs. http://www.census.gov/prod/cen2010/briefs/c2010br-01.pdf. retrieved June 10, 2015

[14] U.S. DOE/EIA, 2014. Environment: State-level energy-related carbon dioxide emissions 2000-2011 U.S. Department of Energy, Energy Information Administration http://www.eia.gov/environment/emissions/state/analysis/