A study of GHG emissions from Canada's fossil-fuel fired electric power systems using the wiki Global Energy Observatory (GEO)

Rajan Gupta and Harihar Shankar
Los Alamos National Lab, Los Alamos, NM 87544

The Global Energy Observatory (http://globalenergyobservatory.org/) is a web-based collaborative tool (a moderated Wiki with built in analysis tools) that aims to provide a one-stop site for detailed unit-by-unit information on global energy systems. The goal of building GEO is to integrate data with analysis tools to facilitate an understanding of the dynamics of change in global energy systems, their networks, emissions and impacts on the environment. We have put equal effort into the development of the framework as in data collection in order that GEO has global reach and use. The intended long-term uses of this system are education, resource planning and optimization, informing policy, and market analysis at the global level.

GEO is an on-going effort, and changes are implemented organically in order to provide the updates as soon as possible. We first briefly describe the organization of the data and basic tools in GEO and then demonstrate the integrated capability of the current version by displaying Canada's Greenhouse Gas (GHG) emissions from fossil-fuel fired generating stations captured from Environment Canada website http://www.ec.gc.ca/pdb/ghg/onlineData/dataSearch_e.cfm. These data are compared with EDGAR (The Emissions Database for Global Atmospheric Research at http://edgar.jrc.ec.europa.eu/index.php) and CARMA (http://www.CARMA.org), two other global emissions databases. We find that data in GEO are more accurate, and claim that the framework is more flexible for collaborative development and real time analysis. To make these data better suited for use in climate and environmental models (dispersal of emissions from point sources) we attempt to geo-locate all infrastructure. Finally, we document in an easy to access summary, the decrease in emission of GHG from fossil-fuel power plants in Canada due to ongoing closure of coal plants, fuel substitution to natural gas and increase in renewable generation. The goal of the GEO framework is to facilitate, for all countries, such studies of the dynamics of change in energy systems and emissions from them.

Organization of the Data in GEO:

GEO is a web interface implemented on top of a relational database. Data for an individual energy related infrastructure are collected into a factsheet. Each factsheet is organized into modules that provide information on different aspects of the project such as location, structural details, annual performance figures, emissions, owner and operator data, comments and references. All infrastructure projects are geospatially referenced using an API to Google Maps and all performance, consumption, emissions, deployment, upgrade and investment data are time referenced.

To facilitate search through the database using GEO's menu based implementation, individual factsheets have been organized hierarchically into four layers: Categories→Types→Countries→States. States (or provinces) are the largest geographical sub-divisions within a country. A brief description of Categories and Types is as follows:
Categories: There are four categories of energy systems labeled (i) Power Plants (Generation Systems); (ii) Fuels and Resources; (iii) Energy Transmission infrastructure; and (iv) Consumers. The tabs in the header line of the webpage connect to dropdown menus that allow viewers to explore the data and tools in each Category. Data in each of these four Categories is further classified by the Type of infrastructure:
Power Generation Systems: Coal, Gas, Geothermal, Hydro, Nuclear, Oil, Solar PV, Solar Thermal, Waste and Wind Turbine electric power generation systems.
Fuels and Resources: Oil and Gas Fields; Coal and Uranium mines; Crude Oil Refineries; Solar and Wind Potential; biomass and water resources; CO_2 sequestration Projects.

Energy Transmission Infrastructure: Oil and Gas Pipelines; Coal, LNG and Oil Ports; Rail, Road and Shipping Links; and the Electric Power Grid.
Consumers: Industry, Population distribution, Economic Activity, and Towns and Cities.

To query and analyze data, the interface provides a menu driven system that follows this hierarchy. Help files are linked to the home page to explain the data fields and facilitate navigation through GEO. Translations of these files to other languages are most convenient using the Google Chrome browser.

Data Collection and Management:

GEO provides the public a framework for data collection, management and analysis in a single unified format. Data are being collected and collated by harvesting (i) open ``official'' databases that exist as fragmented files in heterogeneous (digital and non-digital) formats; (ii) scholarly studies by academia, industry and analysts; and (iii) information volunteered by the public. Our goal is to encourage public participation, i.e., GEO is a moderated wiki. Each factsheet is a web-editable page. To facilitate analysis, individual pages are organized into modules to provide structure and all numerical data are stored and displayed in fields that only accept a single well-defined number.

Finally, to help build and visualize the interconnections between energy systems (local and regional networks), data are cross-referenced (through a user populated module called Associated Infrastructure). This network is displayed using the API to Google Maps when the page is loaded.

By providing a framework that facilitates public participation we aim to address some of the historic challenges of maintaining homogeneity and timeliness of data, its scientific accuracy, completeness, integrity and validity in such global databases. We encourage users to either directly add data to GEO or send it to us for integration.

Status of Tools Integrated with Data in GEO:

Our future efforts will focus on developing a suite of tools to manipulate and analyze data collected in GEO. Current tools include:

Geospatial visualization of data: The tool *"**Map Data in GEO"*** provides an overview of what data exist in GEO and where it is geo-located. To visualize spatial correlations over a geographical region, the user can select either a country for which GEO has data from a displayed menu or a rectangular grid specified by a range of latitude and longitude values. The "show map" button displays the region using Google Maps and an expandable menu of Categories and Types. The viewer can toggle multiple Categories and Types of data stored in GEO and display them (place marks, connecting lines and bounding polygons marking each infrastructure) on the Google Map. The detailed factsheet on an individual project is linked to displays on this map and can be viewed by clicking on the displayed place mark.

Tools within a factsheet: Having loaded a factsheet, data in the Annual Performance module that includes performance indicators, resource utilization and emissions can be graphed versus years. These graphs are used to expose time evolution and correlations in the data. They also facilitate visual consistency checks of the data. In addition, simple algorithmic checks to flag unexpectedly large deviations are run in the background.

Analysis: A large focus of our future effort will be to develop tools to analyze data under each Category and between Categories and Types. Currently, for power plants, the tools allow display of (i) the data for individual units, i.e., the Units module in the factsheet is displayed as a table for all plants in a country or state; (ii) new/cumulative installed capacity versus year of a given Type and for a given country or state;

and (iii) an individual plant's performance versus years as either a table or a figure, along with a projection to a selected future year based on a linear fit to stored data and, for comparison, the median value.

Moderation of data volunteered by the public: The GEO framework incorporates a moderation process that is akin to the peer-review process followed by scientific journals. Any viewer can register and login to edit data. Data volunteered by the public are stored in the database and are visible to the public but do not become part of the default, displayed page (information considered moderated and displayed by default as the editable page). For data to be incorporated and become part of the new moderated version it is reviewed and validated by a subject matter and area experts. The GEO framework has two levels of moderators, i.e., referees and editors. Viewers are expected to provide references to all corrections/additions submitted to facilitate the validation process. The complete history of submissions and edits is maintained in GEO and can be accessed using the "History of Edits" link on the factsheet.

Early Users and Target Audiences:

GEO has been developed to allow any interested person, not just an expert, easy access and visualization of the energy systems. Initially, however, we anticipate the following communities to benefit most from the data and tools in GEO and constitute the early contributors group:
Researchers: By integrating all open-source data into a single database and preserving the timeline we hope to save successive generation of students/researchers from having to rediscover and re-collate the data. The goal is to democratize historic data and facilitate both small and large teams to develop innovative analysis to address the pressing challenge of providing affordable and clean energy to the global population. In addition we hope to motivate researchers and area experts to share volumes of high quality data buried in masters and Ph.D. theses, and utility and analysis companies to share their databases and help enhance the quality and depth of future analysis.
Journalists: Data in GEO can provide a starting point for developing a report on a particular infrastructure project or its role within a network. New findings can be fed back to enhance data in GEO and provide future users with a referenced history of changes to the plant, important landmark events and timelines, responses to regulations and analysis.
Advocacy Groups: GEO would provide a common scientific database around which different stakeholders can build their arguments and analysis.
Utility and Energy Companies: GEO provides a globally accessible database through which companies can showcase their projects and responsibility towards environmental management. Through transparency and verifiable scientific information GEO will highlight highly efficient systems that have smaller environmental footprint and lower GHG gas emissions. Our goal is to identify the best-of-a-kind infrastructures that can serve as examples to emulate and sharing of information on options and opportunities.

Comparison of Data in GEO with EDGAR and CARMA databases:

A growing scientific endeavor is to map the dispersal of emissions from industrial infrastructure to understand impacts on the environment and the climate. The goal of GEO is to provide accurate and up-to-date data on emissions from point sources (such as from stacks in power plants) that can be used as input into transport models for climate change and environmental impact studies. The most exercised part of GEO in terms of data and tools is the Power Plants Category. To evaluate the efficacy of the GEO project, we compare the CO_2 emissions data for Canada with that from EDGAR (the data file EDGAR_v41_CO2_2005_IPCC_1A1_ENE.txt was retrieved from http://edgar.jrc.ec.europa.eu/index.php on 3/2/2011) and CARMA (data downloaded from http://www.CARMA.org on 3/2/2011). The Greenhouse Gas (GHG) data in GEO are copied

from the Environment Canada database that gives facility self-reported estimates for the years 2004-2009. To make the comparison transparent, the complete data and figures used in this analysis are available at http://globalenergyobservatory.org/docs/analysis_results/index.php and here we summarize the main findings:

- The CARMA and EDGAR sources are not accurately geo-located. In fact, we find most of the infrastructure/emissions are not close to the actual location listed and we did not find a pattern in the deviations. The accuracy of data in GEO has been crosschecked using Google Earth, open-source information and official Canadian data.
- To account for the possibility of small deviations in geo-location in EDGAR data, which are provided as total emissions from all industrial sources within a 0.1x0.1 degree cell, we examined the 3x3 block of 0.1x0.1 degree cells centered around the co-ordinates of each plant captured in the GEO database. We find no correlation between the two datasets. In the few cases where some of these nine cells have comparably large CO_2 emissions, we find the total emissions from these nine cells are less than those provided by Environment Canada (data in GEO). The EDGAR values should, if anything, be larger since they are listed as the total emissions from all industrial sources and not just from power plants.
- We matched the CARMA and GEO data using the names of power plants. There is, for most plants, reasonable (within 10 percent) agreement between CARMA estimates and data in GEO for the year 2007.
- Both CARMA and EDGAR databases are static and dated. They do not track changes and therefore do not take into account new capacity additions and/or shutdowns or changes in technology and fuel mix.
- Large differences between the median and GEO emission projections for 2011 should alert the reader to possible major changes in the power plant's use (shutdown, upgrades or reduced use). We have also included plant-by-plant plots of data in GEO in the data file to help viewers visualize the changes.

This comparison highlights that while all three databases are only as good as the quality of data input and frequency and timeliness of additions to it, the advantage of the GEO framework is that it allows public participation for validation, corrections and additions. As demonstrated by the study of emissions from fossil-fuel power plants in Canada, the EDGAR and CARMA data give a dated qualitative picture but are not useful as input for modeling transport of emissions and their impacts on the environment and the climate.

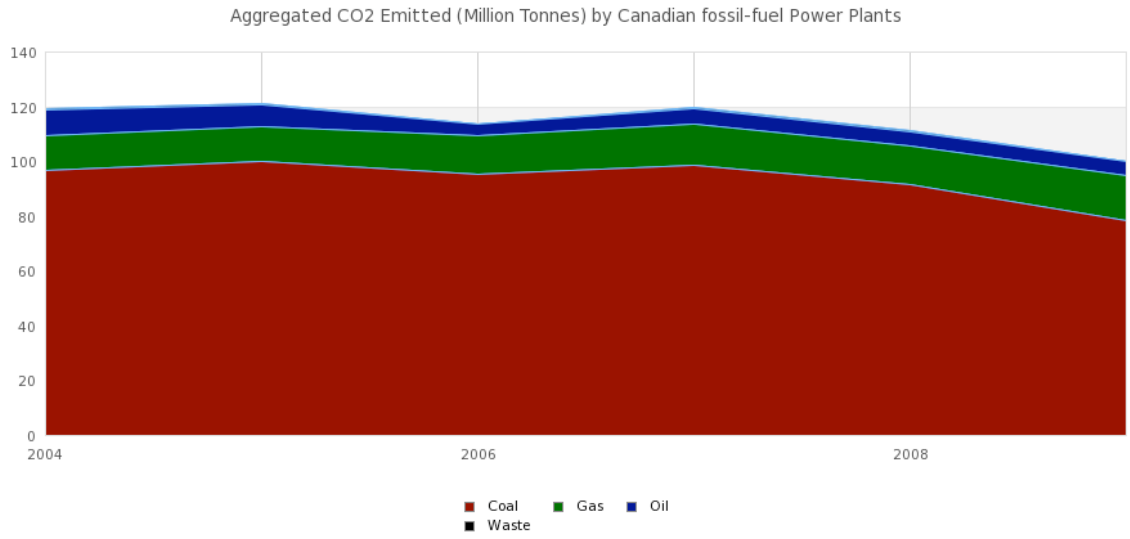

Figure 1: CO_2 emissions data in million tonnes from Environment Canada for coal, gas, oil and waste fired power plants. Data show that between 2004-2009 Canada's emissions have reduced from 121 to 100 million tonnes, mainly due to the retiring of old coal- and oil-fired power plants. The total capacity has increased due to installation of gas turbines and renewable generation, and reactivation of mothballed nuclear reactors.

Finally, in Figure 1, we document Canada's attempts to reduce their CO_2 emissions. The data are aggregated over all the coal, gas, oil and waste-fired power plants in the GEO database. Compared to the highest annual emissions of 121 million tonnes of CO_2 in 2005, the number in 2009 was 100 million tonnes, a reduction by 17 percent. Most of this decrease is due to reduction in use of coal- and oil-fired plants and shutdowns, with Ontario leading the way. There is, on the other hand, an increase of about 3.5 million tonnes from the gas turbine plants, and this number is expected to grow with many more plants coming online in 2010-2012. We note that these estimates do not take into account changes in emissions upstream of the power plants, i.e., in extracting, processing and transporting the various fuels. The bottom line is that while these conclusions are not new, what GEO provides is a transparent view of all the data in one place and the ability to repeat the analysis in real time as and when additional data become available.