Vulcan Project Visualization: A System for Visual Exploration of CO_2 Data

Abstract

Climate Change has emerged as one of the grand global challenges facing humanity. The dominant greenhouse gas contributing to problem, carbon dioxide (CO_2), has a complex cycle through the atmosphere, oceans and biosphere. The combustion of fossil fuels remains the largest source of anthropogenic CO_2 to the atmosphere and its patterns are of central importance to the climate change problem. The quantification of fossil fuel CO_2 was understood only at coarse space and time scales until a recent research effort, Vulcan Project, greatly improved the space/time quantification resulting in data at a resolution of less than 10 km^2/hourly. Using visual tools to examine this new, high resolution data, we can better understand the way that CO_2 is transported within the atmosphere. We have developed visual analytic tools, which allows for easy data manipulation and analysis. Resulting images of the system are used to communicate the data to scientists and non-experts.

Keywords

Visual Analytics, CO_2, Multivariate and Spatio-Temporal Data, Volume Visualization, GIS, Interactive Visualization

Authors

Nathan Andrysco – Department of Computer Science – Purdue University
Kevin Gurney – Department of Earth and Atmospheric Sciences/Department of Agronomy – Purdue University
Bedrich Benes – Department of Computer Graphics Technology – Purdue University
Kathy Corbin – Department of Atmospheric Science, Colorado State University

Introduction

The problem of human-induced climate change has emerged as one of the grand challenges facing humanity. Advances in scientific understanding in the last twenty years have confirmed that climate change is happening, will likely accelerate, and is being driven primarily by rising carbon dioxide (CO_2) levels caused by the combustion of fossil fuels. To stabilize greenhouse gases and avoid further interference with the climate system, CO_2 emissions must decline. However, the current understanding of CO_2 emissions at space/time scales that are meaningful

for carbon cycle science, climate change research, decision-making, mitigation, public understanding, industrial management, and carbon trading is incomplete [Gurney et al., 2007]. Given the profound nature of the climate change problem and the significant trading and infrastructural changes implied by national and international legislation, this limited understanding could cripple the advance of scientific understanding, optimal policy, corporate action, and public engagement. Visualization of the complex space/time dimensions of fossil fuel emissions is a powerful means by which to explore and understand this dominant driver of climate change. This paper describes a visualization system built to better understand the Vulcan Project inventory.

Vulcan

The Vulcan Project [Vulcan 2008] is a NASA/DOE funded effort under the North American Carbon Program (NACP) to quantify North American fossil fuel CO_2 emissions at space and time scales much finer than has been achieved in the past. The Vulcan Project achieved the quantification of the 2002 United States fossil fuel CO_2 emissions down to the level of individual factories, power plants, roadways and sub-county residential areas and placed the complete inventory on a common 10 km^2/hourly grid to facilitate atmospheric modeling. The CO_2 emissions are also subdivided by sector (commercial, residential, industrial, utility, agriculture), by industrial classification (e.g. aluminum smelting, paperboard manufacturing, etc), and fuel.

The original purpose of the Vulcan Project was to aid in quantification of the North American carbon budget, to support inverse estimation of carbon sources and sinks, and to support the demands posed by the launch of the Orbital Carbon Observatory (OCO) [Gurney et al., 2005; Denning et al., 2005; Crisp et al., 2004]. Since the April 2008 release of the Vulcan inventory, it is clear that there are a variety of other uses and user communities interested in the Vulcan results. For example, the detail and scope of the Vulcan CO_2 inventory makes it a valuable tool for educators, policymakers, demographers, and social scientists.

Data

The Vulcan Project produces a large amount of complex, multidimensional data. As such, it holds a variety of meanings for different user communities. For example, atmospheric scientists are interested in the manifestation of surface emissions in the atmosphere and are less-concerned about the processes giving rise to the emissions. They may drive atmospheric transport models to perform inverse estimation of carbon sources and sinks or explore gaps in CO_2 monitoring systems. Policymakers, by contrast, are keenly interested in the emitting processes and what measures can be employed to mitigate emissions. The public may be interested in only a specific spatial domain or a specific sector of emissions.

In all of these endeavors, the information is most effective when it can be analyzed in ways useful to the many potential users interested in CO_2 emissions. Visualization represents a powerfully intuitive means to interpret and understand the Vulcan output. The data produced by Vulcan Project could be visualized through standard volume or GIS visualization techniques. However, the sheer size of the data (over 2 GB per variable for only 4 months of data) requires new techniques in order for the interesting aspects of the data to be effectively communicated.

Emissions

Vulcan Project's data inventory includes significant process-level detail, dividing the U.S. fossil fuel CO_2 emissions into economic sectors and sub-sectors in addition to 23 fuel types. The inventory relies to a great extent on the air quality reporting infrastructure maintained by the United States Environmental Protection Agency (EPA). To meet the air quality standards established by the Clean Air Act, extensive reporting of criteria air pollutants (CAPs) and hazardous air pollutants (HAPs) from nearly every emitting source was made and warehoused by the EPA [USEPA 2005; CFR 2000]. In addition to emissions reporting, a number of other key attributes were submitted, including emission controls, locations, fuel, source classification, and combustion technology. The EPA data was combined with a number of other data and model sources including information on mobile sources, power plants and the U.S. census. The goal is to transform the data, constructed to meet air quality regulations, into a fine-scale fossil fuel CO_2 emissions inventory. Figure 1 shows the broad sectoral breakdown of the Vulcan CO_2 emissions at the "native" resolution of the output (in original non-gridded format). The combination of these data sources comprises the near-complete fossil fuel CO_2 emissions in the U.S. Currently, the only emissions not included from the emissions inventory are from non-road vehicles which includes snowmobiles, trains, tractors, and aircrafts. In addition to the space and time detail, process-level information is retained on all emitting sources including the source classification code (specific industrial combustion devices, sizes, emission controls) and fuel type. To facilitate atmospheric transport modeling and inter-comparison with other independent sources, the CO_2 emissions are placed onto a 10 km x 10 km grid. All point and mobile sources resident within a 10 km x 10 km grid cell were summed while area sources were apportioned via area weighting.
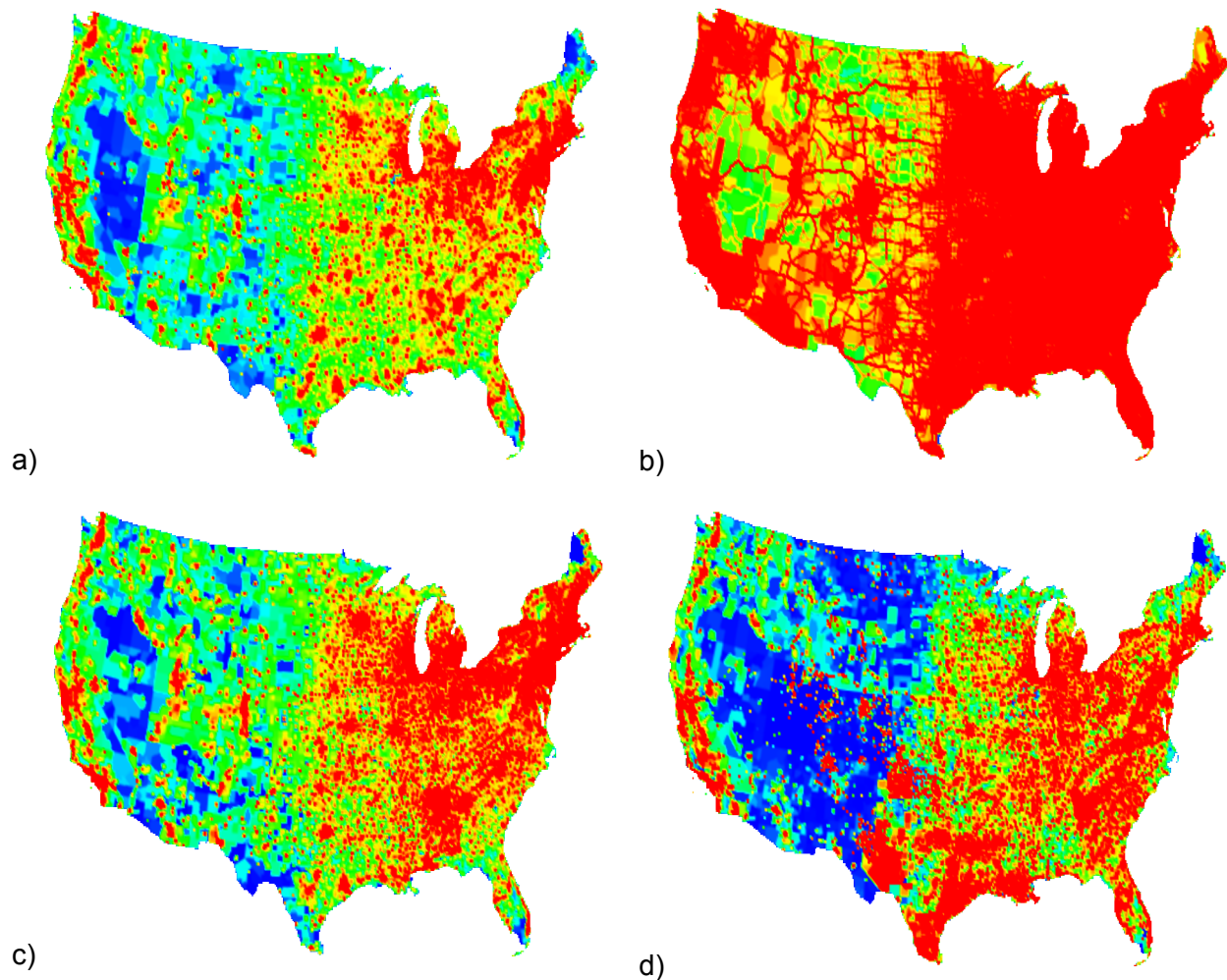
Figure 1: Different CO_2 sources as reported by the Vulcan Project Inventory. a) commercial, b) mobile, c) residential, and d) industrial

The complete 10 km^2/hourly grid constitutes roughly 13 GB of data. Hence, to render each of the five common economic sectors separately (residential, utility, transportation, commercial, industrial, cement) constitutes roughly 80 GB. The native data, prior to regularized gridding, is much more extensive and contains much greater information content such as physical address, fuel type, vehicle type, stack height, and emission controls.

Atmospheric Transport

To better understand the contribution of the fossil fuel emissions to the overall atmospheric concentration of CO_2, the Vulcan emissions can be simulated through the atmosphere of transport models. The emissions inventory of the Vulcan project has been run through the Regional Atmospheric Modeling System (RAMS), which is a mesoscale meteorological model and contains time-dependent equations for velocity, non-dimensional pressure perturbation, potential temperature, ice, water, and cloud microphysics. A significant feature of the model is the incorporation of a telescoping nested-grid capability. RAMS has recently been coupled with the Simple Biosphere Model (SiB), a landsurface parametrization scheme originally used to compute biophysical exchanges in climate models. The input variables provided by RAMS to SiB are updated every minute during the simulation time. SiB provides RAMS with fluxes of heat, moisture, momentum and CO_2, as well as upwelling radiation. The coupled SiB-RAMS model has been used to study PBL-scale interactions among carbon fluxes, turbulence, and CO_2 mixing ratio and regional-scale controls on CO_2 variations. The Vulcan emissions are re-gridded to the 40 km RAMS grid and temporally discretized to the three minute time step of the transport model. The 40 km North American ETA reanalyzed winds are used to drive the atmospheric transport. To simulate four contiguous months, the Vulcan-RAMS analysis required roughly one week of computation time on a 50 node Linux cluster. Figure 2 shows the annual mean CO_2 contribution difference between the Vulcan inventory and the previous state-of-the-art inventory [Andres et al. 1996; Brenkert et al., 1998] at 30 m height above the surface.
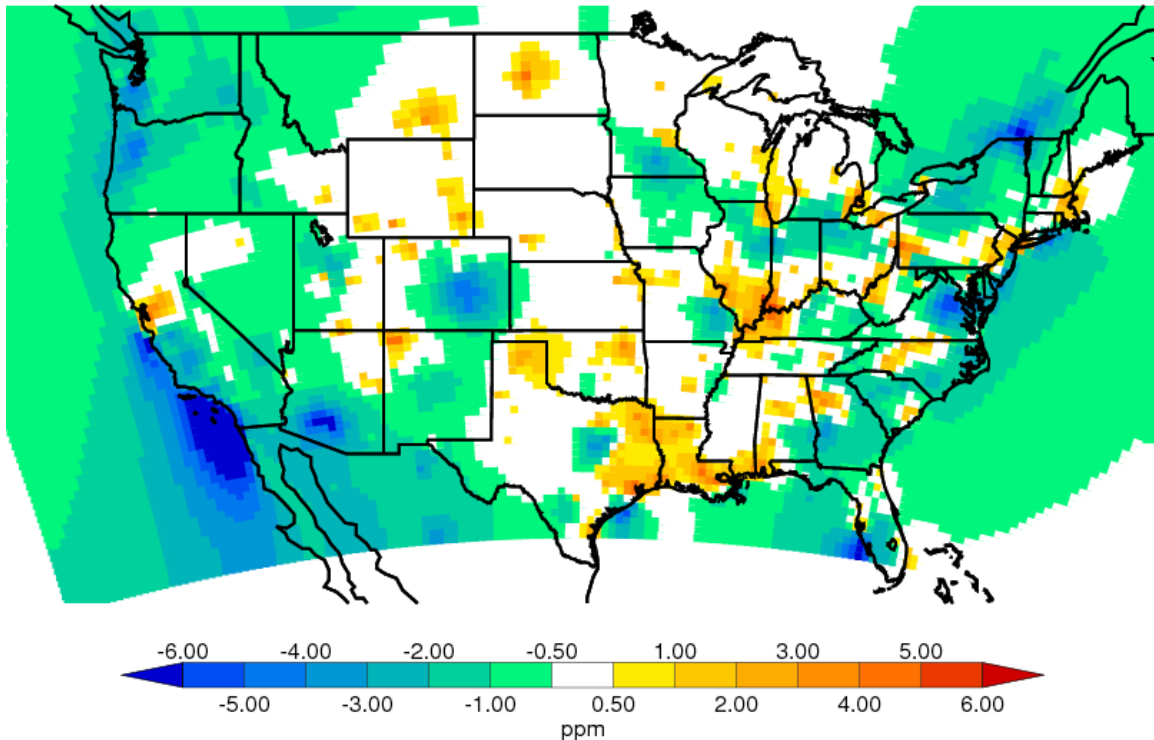
Figure 2: Annual mean 30 m difference between simulated Vulcan and Andres et al., (1996) fossil fuel CO_2 emissions inventories (Vulcan – Andres).

Visualizations

The data produced by the Vulcan Project interests different user communities in a variety of ways. For example, atmospheric scientists are interested in using the visualizations to extract important features resident in the data. In particular, they need to explore not only the sources of CO_2 but also the transport within the Earth's atmosphere. The general viewing public is more interested in the sources of CO_2, particularly their "own" CO_2 footprint. In order to facilitate these needs, background political and elevation maps are provided to give the viewer geographic reference points. One of the goals of our visualization was to help people develop a deeper understanding of the sources of CO_2 at multiple scales. Visualization can clarify the contribution of the average person's CO_2 emissions not only in their own neighborhood, but the extent of that impact across the country and over U.S. borders and across oceans. Policymakers are another important group of users. Policymakers do not necessarily have the scientific background of the atmospheric scientist, but they typically have a deeper understanding than the general public. Their primary interest is to use the visualizations to understand the extent

and nature of current emissions in order to construct mitigation measures, build a carbon trading market, and communicate the need for new policy.

Since the emissions data is of great interest to both the general public and policymakers, the visualization system must allow the data to be viewed in different ways. The emissions data is a set of 2D data points located on the surface of the United States. The user can view the pre-discretized data, such as the mobile sources overlaid on top of major U.S. roadways, traffic by each state, etc. The transport of CO_2 can also be viewed using area visualization techniques, which is traditionally how researchers have looked at their data in the past. Since the atmosphere exhibits extremely complex flow, the use of 2D slices have been the easiest way for them to isolate interesting events in the data. One technique provided is achieved by using a color map and blending/shading capabilities to create a composed image of CO_2 values. To visualize areas with higher than critical values of CO_2 and their evolution vertically in the atmosphere, isolines are generated, as depicted in Figure 3.
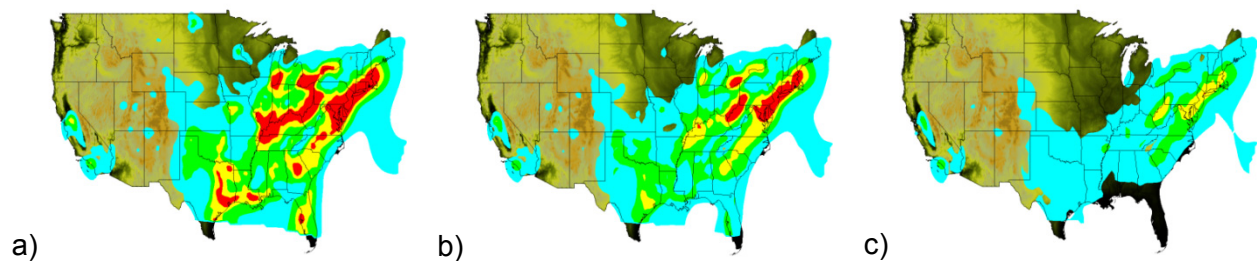

a)   b)   c)

Figure 3: CO_2 concentrations at different heights in the atmosphere. a) 35 m b) 75 m c) 180 m. CO_2 concentrations reflect the emissions at the surface, but become dispersed in higher altitudes of the atmospheric layer

A number of interesting features are evident in the surface 2D visualizations. Most notable is the diurnal cycle, or the daily pattern, of the emissions field itself. This surface concentration is a close reflection of the surface emissions but modified by the diurnal cycle of the boundary layer transport, a quasi-compressible layer of air near the surface of the earth. If emissions were constant over a 24 hour cycle, the surface layer CO_2 concentrations would still exhibit a diurnal cycle due to the direct compression of the boundary layer in the night and early morning. The emissions are not constant throughout the day but also follow day/night patterns. The mobile sector has a "rush hour" structure with heightened emissions during local morning and local evening hours and little emissions during the night. Though power production emissions have a complicated structure due to the tradeoff power companies make with base-load and peak power demand, they do show heightened emissions during daylight hours due to commerce and increased demands for heating and cooling. The combination of the diurnal cycle for the

boundary layer with the diurnal cycle of emissions generally results in greater concentrations during the day and less concentrations at night. The peak concentrations occur in the early morning hours when the compression of the boundary combines with a rapid increase in emissions at the surface. Viewing this cycle as an animation has lead journalists to call it the "breath of the nation" [Revkin 2008].

Another feature immediately evident in the surface visualizations are the broad maxima over populated or industrially-intensive locations. This is driven by the surface sources. Noticeable are the populated West coast centers of Los Angeles/San Diego, San Francisco Bay Area, Portland and Seattle. Moving eastward, one notices the Salt Lake City urban corridor and the Denver front range, followed by the Houston and Dallas areas. Finally, the population centers of the upper Midwest, the Ohio Valley and the Boston-New York-Washington mega-urban corridor.
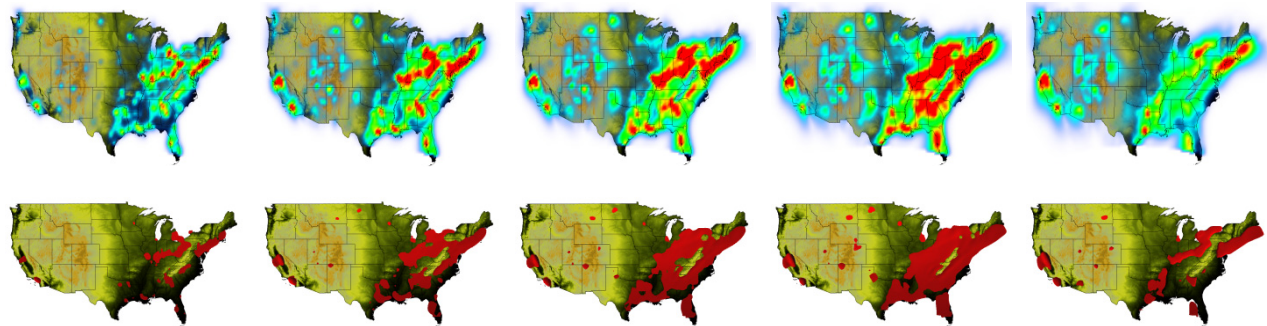

Figure 4: A comparison between the emission (top) values and concentrations (bottom) over a number of time steps during May (the beginning of the Summer data set). Concentrations strongly reflect the emissions during the early time period. Later time steps show the accumulation of emissions as the concentrations become larger and propagate higher into the atmosphere.

The 3D isosurface rendering improves insight into phenomena such as CO_2 transportation and the way they are affected by weather fronts, which had previously been difficult to extract using the off-the-shelf visualization methods. One such example is the summertime transport of air in Southern California southward over the tropical Pacific and Mexico, which is clearly noticeable in the 3D visualizations. The transport over much of the upper tier of the continental portion of the U.S. is also readily noticeable (see accompanied video). Another interesting phenomenon which is evident is the geotrophic flow from west to east. An example of this is CO_2 moving off of the East coast out over the North Atlantic. The summertime influence of Gulf air is evident as well with some of the CO_2 washing out over the gulf, most likely due to lower-level mixing in summer thunderstorm activity. This same thunderstorm activity is also likely responsible for the

middle- and late-summer elevated parcels of CO_2 rich air evident in the isosurface visualization. These detached parcels are likely the result of surface air transported rapidly upward to thunderstorm outflow regions, where they remain in coherent form. Frontal systems are evident in the upper tier of the continental U.S. as large wedge-shaped features likely denoting fronts of warm air "climbing" up low-level wedges of polar air outbreaks. Some of the cyclonic motion can be discerned as well.

Visualizing only quantitative values of CO_2 concentrations at different atmospheric layers, without any regard to latitude/longitude position, highlights another insight into the CO_2 transport. A histogram of CO_2 concentrations is generated by projecting CO_2 concentrations over a predefined area to horizontal line segments. Which line segment a concentration is mapped to is determined by the position of the concentration in the atmosphere. CO_2 concentrations located near the surface are mapped to the lower horizontal lines whereas those concentrations at the upper atmospheric layers are placed on the upper horizontal lines. The *x* position is determined by the CO_2 concentration, with low concentrations mapped to the left. Each latitude/longitude point has multiple atmospheric layers and the layers of each projected latitude/longitude are connected together with a line. If viewing just a single projected geographic point, the line created will show how the CO_2 concentration changes starting from the layer nearest to the surface and up through the atmosphere. The lines are also colored based on the density of points with similar CO_2 concentrations. Red indicates there are many similar CO_2 concentrations at the given height level in the atmosphere, with blue indicating few CO_2 concentrations. To emphasize the point, color indicates the density of CO_2 concentrations and the location of each point on the horizontal line indicates the actual value of the concentration. Figure 5 is an example of this visualization for the entire United States. The majority of CO_2 concentrations are at the low end of the spectrum, as indicated by the large section of red on the left-hand side.
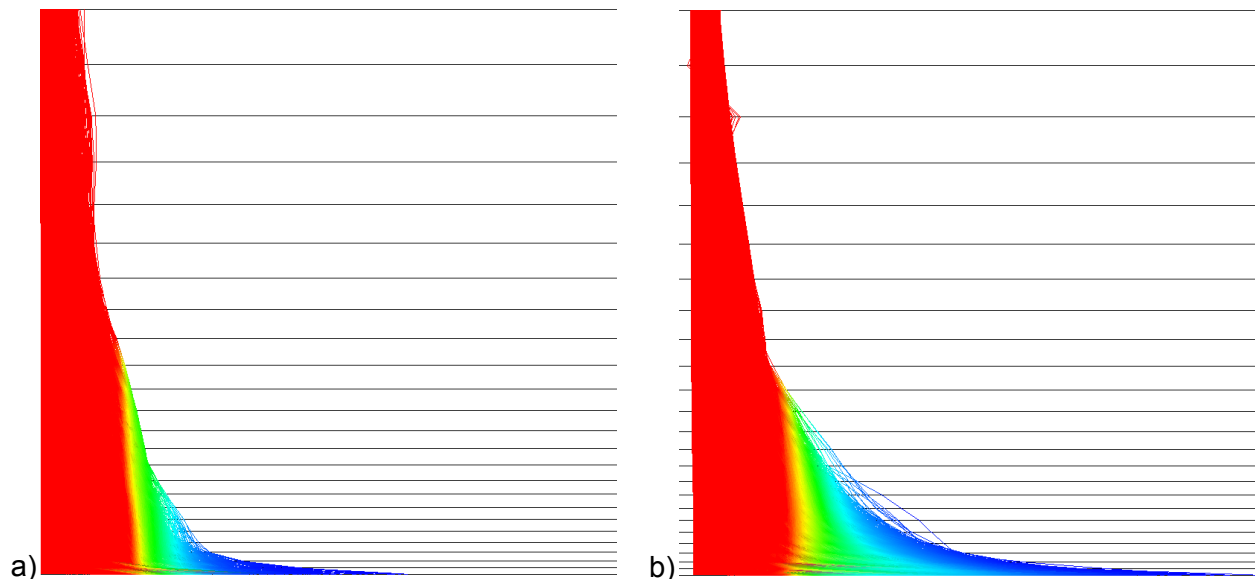
Figure 5: a) A histogram for the entire U.S. in early May. b) A histogram for late August. Each histogram's bottom horizontal line is for CO_2 concentrations nearest to the surface (~35 m) and the upper line maps concentrations at ~3,400 m. Notice how the CO_2 values have shifted to the right. This indicates that there has been an increase in overall CO_2 concentrations over the U.S.

The histogram visualizations show a number of interesting features that are not apparent in the other visualizations. Near the surface, the spread of concentration is evidence of both the spectrum of source function strength and perhaps spatial variability in the boundary layer height. This spread shows the diurnal cycle as well. The convergence of values just above the surface likely denotes the boundary layer lid. Above this lid, there appears to be irregular spreading and compression of the vertical profile traces. This is likely due to the penetration of synoptic events that advect CO_2 rich air above the source locations downwind which weakens the horizontal gradients by eliminating the influence of the surface source function.

Even though it is believed that the RGB color scale is not the best choice in most situations [Borland 2007] the familiarity that people have with it was crucial when communicating CO_2 values. The use of cool colors to indicate low values and warm colors to indicate high values helps communicate the need for environmental change. Psychologically seeing red illicits an emotional reaction associated with danger. Blue and other cool colors, give the impression of safety. When people see a large portion of the country go from blue to red in the visualizations, it is immediately understood that "safe" CO_2 levels have been replaced by "dangerous" levels.

Impact

Release of the first Vulcan fossil fuel CO_2 emissions inventory occurred in April 2008 and its subsequent release to the public drew enormous attention from the media, scientists, and the public, which most likely reflects the importance of the project. In addition to the establishment of the Vulcan website [Vulcan 2008] a video of various aspects of atmospheric transport was released onto YouTube [YouTube 2008]. The YouTube video got an incredible 100,000 views in just two days and was close to 200,000 views in July 2008. The Purdue press release [Purdue 2008] drew the attention of national media outlets such as NY Times (Dot Earth series: dotearth.blogs.nytimes.com), the Boston Globe (full page Sunday edition), Reuters, Scientific American (a video story), Wired Magazine, National Public Radio, and the Discovery Channel. Moreover, the press release appeared in over 150 online news sources and countless regional print news outlets, local TV, and radio shows. The story in Wired Magazine drove so much traffic and blogging, the editors requested additional analysis.

The unexpected public interest was so intense in the first 48 hours of release that the Vulcan website host server was brought down by the volume of traffic to the site. The release has generated a flood of email from nearly every corner of the globe. In the first week the Vulcan video was #3 in US, #6 in Ireland, #17 in New Zealand, #5 in Canada and many other countries in "Science and Technology" category of YouTube. Google has donated an engineer to embed the Vulcan results into Google Earth. Senator Richard Lugar opened recent Senate Foreign Relations Committee testimony with praise of the Vulcan Project and has written a "dear colleagues" letter in the Senate highlighting the importance of Vulcan to the legislative work on climate change policy. Carbon cycle scientists have begun utilizing the Vulcan inventory in inverse work, forward integrations, and tracer studies while climate scientists have begun exploring how energy-driven CO_2 emissions may respond to increasingly frequent and severe weather events.

Future Goals

As a result of the success of the Vulcan Project a major new initiative, is currently underway - the Hestia Project [Hestia 2008] - which will quantify greenhouse gas emissions for the entire planet at a building scale with complete underlying driving processes (driving decisions, heating decisions, etc).

One of the biggest challenges when visualizing the data was obtaining interactive rates over multiple time steps. The time it took to execute the geometry generation of the visualizations was limited by memory access time. Due to the dataset size, loading all of the data into main memory was not always an option so an out-of-core visualization mechanism was implemented. If only the subset of data was loaded for each visualization, there was still a bottleneck due to size of data per time step. The next step is to explore methods that will facilitate interactivity.

Bibliography

- Andres, R.J. Marland, G., Fung, I. and E. Matthews (1996), A 1 x 1 distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950-1990, *GBC*, **10** (3), 419-429.
- Borland, D.; Taylor, R.M., "Rainbow Color Map (Still) Considered Harmful," *Computer Graphics and Applications, IEEE* , vol.27, no.2, pp.14-17, March-April 2007
- Brenkert, A.L. Carbon dioxide emission estimates from fossil-fuel burning, hydraulic cement production, and gas flaring for 1995 on a one degree grid cell basis, (http://cdiac.esd.ornl.gov/ndps/ndp058a.html), 1998.
- Code of Federal Regulations (2000), Protection of Environment, Environmental Protection Agency, 40 CFR Part 75, Continuous Emission Monitoring, Code of Federal Regulations, Revised as of July 1, 2000.
- Crisp, D., et al., (2004), The Orbiting Carbon Observatory (OCO) mission, *Advances in Space Research*, **34**, (4), 700-709.
- Denning, A.S. et al., (2005), Science Implementation Strategy for the North American Carbon Program. Report of the NACP Implementation Strategy Group of the U.S. Carbon Cycle Interagency Working Group. Washington, DC: U.S. Carbon Cycle Science Program, 68 pp.
- Gurney, K.R., et al., (2005), "Sensitivity of Atmospheric CO_2 Inversion to Seasonal and Interannual Variations in Fossil Fuel Emissions," *J. Geophys. Res.* **110** (D10), 10308-10321.
- Gurney, K.R. et al. (2007), "Research needs for process-driven, finely resolved fossil fuel carbon dioxide emissions," *EOS*, December 4.
- "Hestia Project". 2008. Department of Atmospheric Science. Purdue University. (http://www.purdue.edu/climate/hestia/)
- Purdue University. "'Revolutionary' CO2 maps zoom in on greenhouse gas sources." April 7, 2007. Press Release. (http://news.uns.purdue.edu/x/2008a/080407GurneyVulcan.html)
- Revkin, Andrew C. "Breath of a Nation — Animated CO2 Map." April 7, 2008. Dot Earth. New York Times. (http://dotearth.blogs.nytimes.com/2008/04/07/breath-of-a-nation-animated-co2-map/)
- USEPA (2005), Emissions Inventory Guidance for Implementation of Ozone and Particulate Matter National Ambient Air Quality Standards (NAAQS) and Regional Hazw Regulations, Emissions Inventory Group, Emissions, Monitoring and Analysis Division, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, EPA-454/R-05-001, August 2005.
- "Vulcan Project, The." 2008. Department of Atmospheric Science. Purdue University. (http://www.purdue.edu/eas/carbon/vulcan/research.html)
- YouTube. "'Revolutionary' CO2 maps zoom in on greenhouse gas sources." (http://www.youtube.com/watch?v=eJpj8UUMTaI)