Editor: Mike Potel
www.wildcrest.com

Visual Exploration of the Vulcan CO_2 Data

Nathan Andrysco, Kevin Robert Gurney, and Bedřich Beneš, ***Purdue University***

Kathy Corbin, ***Colorado State University***

The problem of human-induced climate change has emerged as one of the grand challenges facing humanity. Advances in scientific understanding in the past 20 years have confirmed the fact of climate change and its likely acceleration, driven primarily by rising carbon dioxide (CO_2) levels—mostly from fossil-fuel combustion. We must reduce CO_2 emissions to stabilize greenhouse gases and avoid further interference with the climate system. However, we currently have limited 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.[1] 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 for exploring and understanding this dominant climate-change driver. We've developed a visualization system to help explore and understand CO_2 emission data collected by the US Vulcan Project, a government-funded project to quantify North American fossil-fuel CO_2 emissions at much finer space and time scales than have been achieved in the past.

The Vulcan Project

NASA and the US Department of Energy (DoE) are funding the Vulcan Project (www.purdue.edu/eas/carbon/vulcan/research.html) under the North American Carbon Program (NACP). The project quantified the 2002 US fossil-fuel CO_2 emissions down to the level of individual factories, power plants, roadways, and subcounty residential areas and placed the complete inventory on a common hourly, 10 km^2 grid to facilitate atmospheric modeling. This resolution is much finer than prior estimates that resolved to only the state and monthly levels. The project also subdivided CO_2 emissions by sector (commercial, residential, industrial, utility, agriculture), industrial classification (aluminum smelting, paperboard manufacturing, and so on), and fuel.

The Vulcan Project originally aimed to help quantify the North American carbon budget and to support the demands posed by the launch of the Orbital Carbon Observatory.[2–4] Since the April 2008 release of the Vulcan inventory, various other uses and user communities have emerged. The inventory's detail and scope make it a valuable tool for educators, policymakers, demographers, and social scientists. Because the project data is large, complex, and multidimensional, it holds different meanings for different user communities. For example, atmospheric scientists are interested in how surface emissions manifest themselves in the atmosphere and are less concerned about the processes giving rise to the emissions. Policymakers, in contrast, are keenly interested in the emitting processes and ways to mitigate emissions. The public might be interested in only a specific spatial domain or a source, such as residences or vehicles.

In all these endeavors, the information is most effective when users can analyze it in ways specific to their interests. Visualization offers an intuitive means to interpret and understand the Vulcan output. Researchers can use standard volume or GIS techniques to visualize the data. However, the sheer size of the data (over 2 Gbytes per variable for only four months) requires specialized techniques to effectively communicate its interesting aspects.

CO_2 Emissions

The Vulcan data inventory includes significant process-level detail, dividing US fossil-fuel CO_2 emissions into economic sectors and subsectors in addition to 23 fuel types. The inventory relies on the air-quality reporting infrastructure maintained by the US Environmental Protection Agency (EPA). To meet the air-quality standards established by the Clean Air Act, the EPA established extensive requirements for reporting criteria air pollutants (CAPs) and hazardous air pollutants (HAPs) from nearly every emitting source and

 Published by the IEEE Computer Society

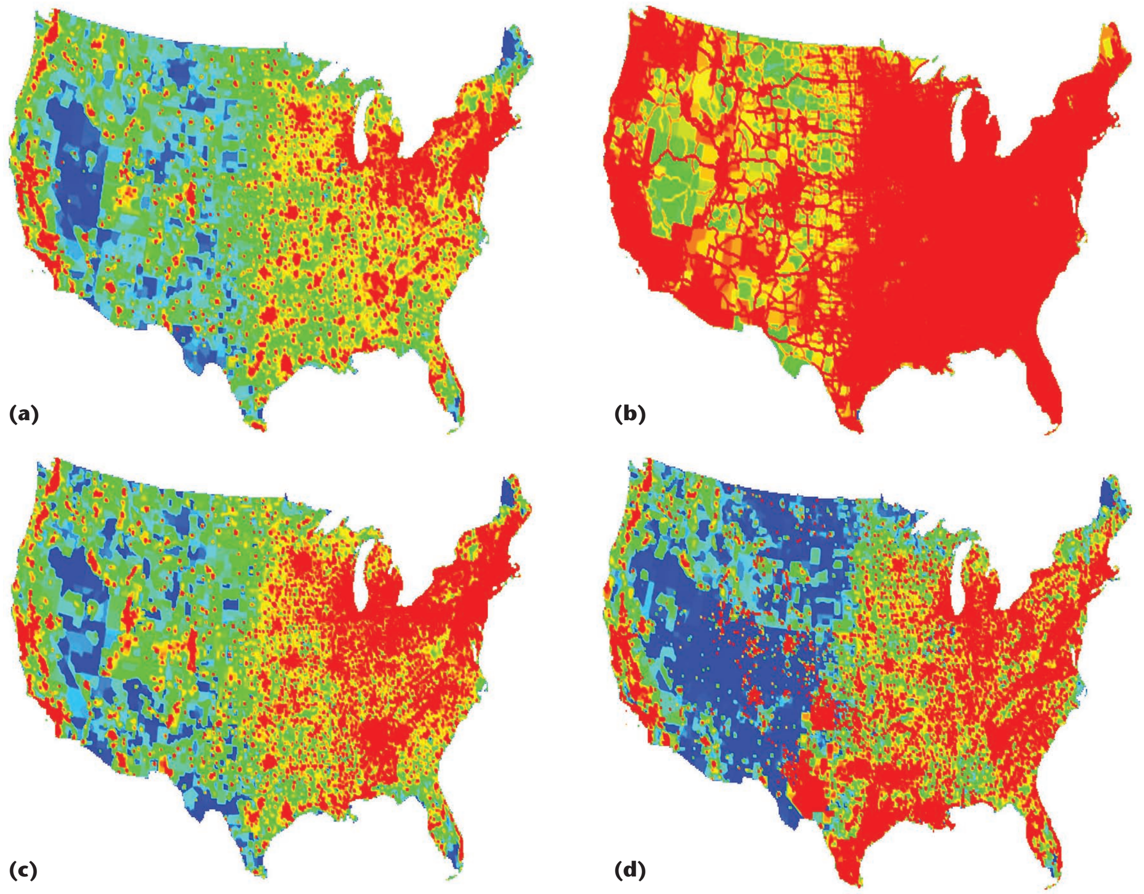

Figure 1. Different CO_2 sources as reported by the Vulcan Project inventory: (a) commercial, (b) mobile, (c) residential, and (d) industrial. The color scale is the RGB standard where blue is zero emissions and red is 625 kilotons of carbon/year. The scale is linear, so green is 312.5. The predominance of red in mobile CO_2 sources indicates its dominance compared to the other sources.

warehoused the resulting EPA data (www.epa.gov/ttn/chief/eidocs/eiguid/index.html and http://vlex.com/source/1089/toc/01.03.63). Besides emissions reporting, local environmental managers submitted other key attributes, including emission controls, locations, fuel, source classification, and combustion technology.

The Vulcan Project combined the EPA data with other data and model sources, including information on mobile sources, power plants, and the US census. The goal is to transform data constructed to meet air-quality regulations into a fine-scale fossil-fuel CO_2 emissions inventory. Figure 1 shows the broad sector breakdown of the Vulcan CO_2 emissions at the output's "native" resolution—that is, the original nongridded format. The combined data sources comprise practically all US fossil-fuel CO_2 emissions. The only emissions not included from the emissions inventory are from nonroad vehicles, such as snowmobiles, trains, tractors, and aircraft. Besides the space and time detail, the data includes process-level information on all emitting sources, including the source classification code (specific industrial combustion devices, sizes, and emission controls) and fuel type. Placing the CO_2 emissions on a 10 km $\times$ 10 km grid facilitates atmospheric-transport modeling and comparison with independent sources. All point and mobile sources in a grid cell were summed, while area sources were apportioned via area weighting.

The complete hourly, 10 km^2 grid constitutes roughly 13 Gbytes of data. So, rendering each of five common economic sectors separately involves roughly 80 Gbytes. The native data, before regularized gridding, is much more extensive and contains information such as physical address, fuel type, vehicle type, stack height, and emission controls.

Atmospheric Transport

To better understand the fossil-fuel emissions' contribution to the overall atmospheric concentration of CO_2, the Vulcan Project simulated its emissions data through atmospheric-transport model called the Regional Atmospheric Modeling System (RAMS), developed at Colorado State University (http://rams.atmos.colostate.edu). RAMS is a mesoscale meteorological model that solves time-dependent equations for velocity, nondimensional pressure perturbation, ice-liquid water potential temperature, total water mixing ratio, and cloud microphysics. The Vulcan Project regridded its emissions to the 40 km RAMS grid and temporally discretized them to the transport model's three-minute time step. The National Center for Environmental Protection's Eta model is used to reanalyze the 40-km North American winds, which then drive the atmospheric transport.

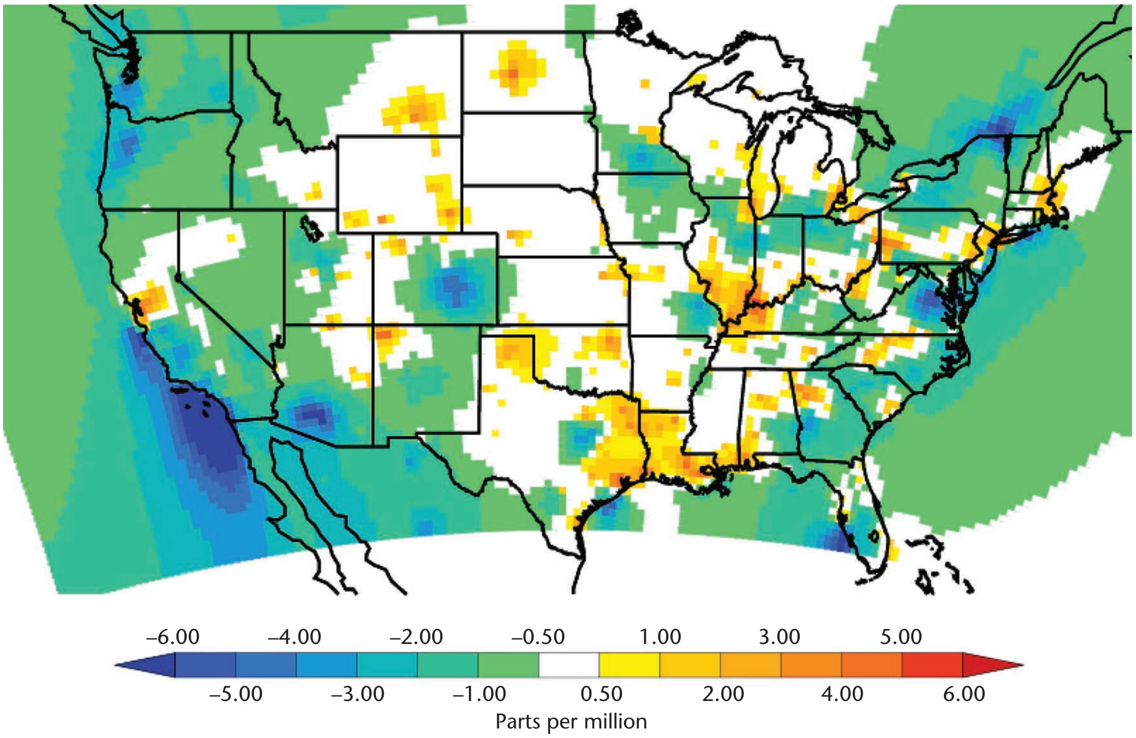

Figure 2. The annual mean CO_2 contribution difference between the simulated Vulcan data and the previous state-of-the-art data,[5] for 30 m above the Earth's surface. For instance, the Vulcan data for the Northwest states of Washington and Oregon shows less CO_2 being produced than the earlier model showed.

To simulate four contiguous months, the Vulcan-RAMS analysis required roughly one week of computation on a 50-node Linux cluster. Much of the computational demand is devoted to solving 3D equations that describe the motion of the Earth's atmosphere. Figure 2 shows the annual mean CO_2 contribution difference between the Vulcan inventory and the previous state-of-the-art inventory, for 30 m above the surface.[5,6]

Visualizations

Among the user communities for Vulcan data, atmospheric scientists need to explore not only the sources of CO_2 but also its transport in the Earth's atmosphere.

The general public is more interested in CO_2 sources, particularly individual CO_2 footprints. Background political and elevation maps give viewers geographic reference points to facilitate their needs. One of our visualization goals was to help people deepen their understanding of the CO_2 sources at multiple scales. Visualization can clarify the contribution of the average person's CO_2 emissions not only in his or her own neighborhood but also across the country, over US borders, and across oceans.

Policymakers are another important user group. They don't necessarily have an atmospheric scientist's background, but they typically have a deeper understanding than the general public. Their primary interest is to use visualizations to understand the extent and nature of current emissions to help construct mitigation measures, build a carbon-trading market, and communicate new policy needs.

The Vulcan emissions data is a set of 2D data points located on the surface of the US. Users can view the prediscretized data, such as data from mobile sources, overlaid on top of major US roadways, traffic patterns for each state, and so on. They can also view CO_2 transport using area visualization techniques, which is how researchers have traditionally looked at their data. The atmosphere exhibits extremely complex flow, so researchers have found 2D slices to be the easiest way to isolate interesting events in the data. One technique uses a color map and blending/shading capabilities to compose an image of CO_2 values. To visualize areas with higher-than-critical CO_2 values and their evolution vertically in the atmosphere, we generate isolines (see Figure 3).

Several interesting features are evident in the surface 2D visualizations. Most notable is the diurnal cycle, or daily pattern, of the emissions field itself (see Figure 4a). This surface concentration closely reflects the surface emissions but is modified by the diurnal cycle of the *boundary-layer transport*, a quasicompressible layer of air near the Earth's surface (see Figure 4b). If emissions were constant over a 24-hour cycle, the surface-layer CO_2 concentrations would still exhibit a diurnal cycle because of the boundary layer's direct compression in the night and early morning. The emissions aren't constant throughout the day but also follow day/night patterns. The mobile sector has a "rush hour" structure, with heightened emissions

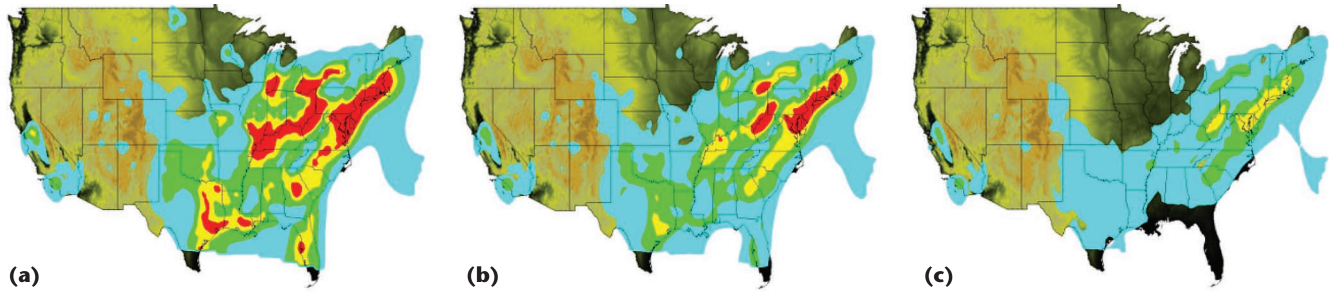

Figure 3. CO_2 concentrations at different heights in the atmosphere: (a) 35 m, (b) 75 m, and (c) 180 m. CO_2 concentrations reflect emissions at the surface but become dispersed at higher altitudes of the atmospheric layer.

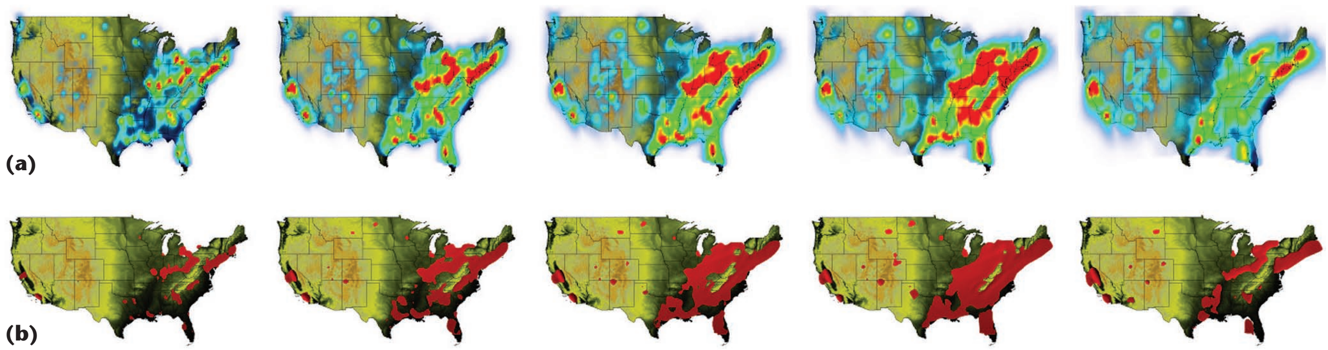

Figure 4. A comparison between the emission (a) values and (b) concentrations over five time steps during May (the beginning of the summer data set). Concentrations strongly reflect the emissions during the early time period. Later time steps show the emissions accumulating as the concentrations become larger and propagate higher into the atmosphere.

during local morning and local evening hours and low emissions at night.

Although power-production emissions have a complicated structure due to the trade-off 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 boundary-layer and emission diurnal cycles generally results in greater concentrations during the day and lower concentrations at night. The peak concentrations occur in the early morning hours when the boundary compression combines with a rapid increase in emissions at the surface. Viewing an animation of this cycle has led journalists to call it the "breath of the nation" (http://dotearth.blogs.nytimes.com/2008/04/07/breath-of-a-nation-animated-co2-map).

The broad maxima over populated or industrially intensive locations is another feature immediately evident in the surface visualizations. Surface sources drive this phenomenon. They include the populated West Coast centers of Los Angeles/San Diego, the San Francisco Bay area, Portland, and Seattle. Moving east, the Salt Lake City urban corridor and Denver front range appear, followed by the Houston and Dallas areas. Finally, we see the population centers of the upper Midwest, the Ohio Valley, and the Boston-New York-Washington mega-urban corridor.

A 3D isosurface rendering improves insight into phenomena such as CO_2 transport and the way weather fronts affect them. These phenomena had previously been difficult to extract using off-the-shelf visualization methods. One such example is the summertime transport of air in Southern California southward over the tropical Pacific Coast and Mexico. The transport over much of the northern continental US is also readily noticeable. Also evident is the geotrophic flow from west to east. An example is CO_2 moving off the East Coast out over the North Atlantic. The summertime influence of Gulf of Mexico air is also evident with some of the CO_2 washing out over the gulf, most likely caused by lower atmospheric-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 likely result from surface air transported rapidly upward to thunderstorm outflow regions, where they remain in coherent form. Frontal systems are evident in the northern continental US as large wedge-shaped features likely denoting warm-air fronts climbing up low-level wedges of polar-air outbreaks. Some of the cyclonic motion is also noticeable.

To give further insight into CO_2 transport, we can generate a histogram of CO_2 concentrations by projecting them over a predefined area to horizontal

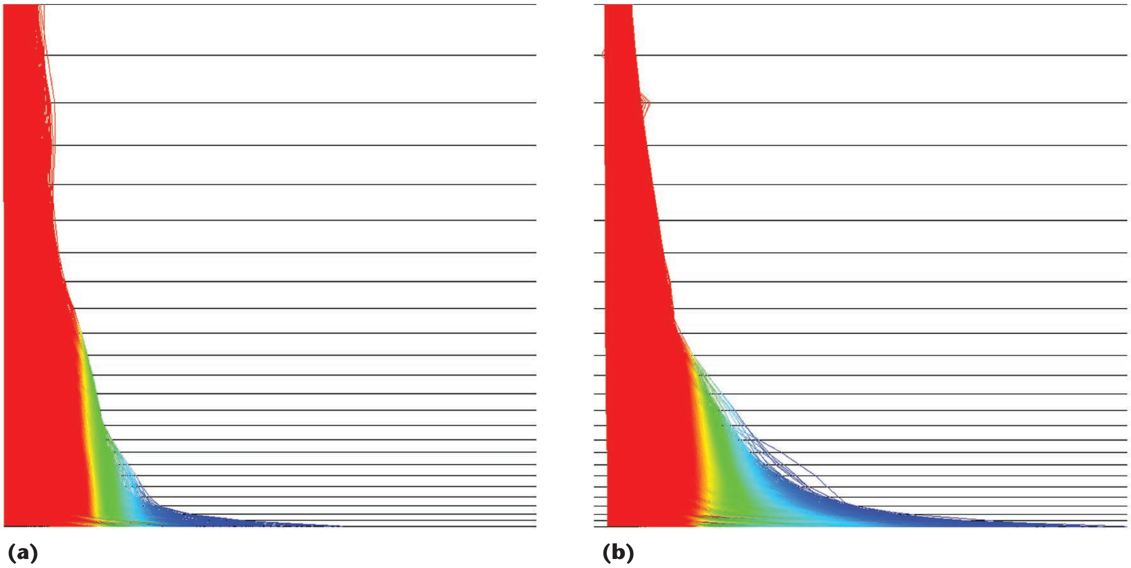

Figure 5. Histograms for the entire US: (a) early May and (b) late August. Each histogram's bottom horizontal line is for CO_2 concentrations nearest the surface (~35 m); the upper line maps concentrations at ~3,400 m. The CO_2 values shift to the right, indicating an increase in overall CO_2 concentrations over the US.

line segments. The line segment to which a concentration is mapped is determined by its position in the atmosphere. CO_2 concentrations near the surface are mapped to the lower horizontal lines, whereas concentrations at the upper atmospheric layers are placed on the upper horizontal lines. The *x*-axis position is determined by the CO_2 concentration, with low concentrations mapped to the left. Each latitude-longitude point has multiple atmospheric layers that are connected by a line. Viewing just a single projected geographic point will create a line that shows how the CO_2 concentration changes, starting from the layer nearest the surface and up through the atmosphere.

The lines are also colored according to the density of points with similar CO_2 concentrations. Red indicates many similar CO_2 concentrations at the given height in the atmosphere, with blue indicating few CO_2 concentrations. Each point's location on the horizontal line indicates the concentration's actual value.

Figure 5 is an example of this visualization for the entire US. Most CO_2 concentrations are at the low end of the spectrum, as indicated by the large section of red on the left.

The histogram visualizations show several unique features. Near the surface, you can see a wide spectrum of CO_2 concentrations. As you move up in the atmosphere, you can see that lower CO_2 concentrations dominate. 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. The likely cause is the penetration of synoptic events that advect CO_2-rich air above the source locations downwind, which weakens the horizontal gradients by eliminating the surface source function's influence.

Even though most researchers don't believe that RGB is the best color scale in most situations,[7] its familiarity was crucial in 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. Seeing red elicits 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, they immediately understand that safe CO_2 levels have degraded to dangerous levels.

Project Data Impact

The Vulcan Project released its first fossil-fuel CO_2 emissions inventory in April 2008, and its subsequent public release drew enormous attention from the media, scientists, and the public. In addition to establishing the Vulcan Web site, the project released a video of various aspects of atmospheric transport to YouTube (www.youtube.com/watch?v=eJpj8UUMTaI). 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 (http://news.uns.purdue.edu/x/2008a/080407GurneyVulcan.html) drew the attention of national media outlets such as the *New York Times* (as part of the Dot Earth series), the *Boston Globe* (a full-page story in the Sunday edition), Reuters, *Scientific American* (a video story), *Wired* magazine, National Public Radio, and the Discovery Channel. Moreover, the press release appeared in more than 150 online news sources and countless regional print news outlets, local TV stations, and radio shows. The *Wired* story drove so much traffic and blogging that the editors requested additional analysis.

The intense public interest in the first 48 hours of release was unexpected, and traffic volume brought the Vulcan site's host server down. The release has generated a flood of email from around the globe. In the first week, the Vulcan video was the third-most-popular video in the US, sixth in Ireland, 17th in New Zealand, and fifth in Canada and many other countries in YouTube's Science and Technology category. 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 for the Vulcan Project and has written a letter to his Senate colleagues highlighting its importance to legislative work on climate-change policy. Carbon-cycle scientists have begun using the Vulcan inventory in carbon budget studies, transport modeling, and chemical tracer studies, while climate scientists have begun exploring how energy-driven CO_2 emissions respond to increasingly frequent and severe weather events.

As a result of the Vulcan Project's success, a major new initiative is underway. The Hestia Project (www.purdue.edu/climate/hestia) will quantify greenhouse gas emissions for the entire planet at a building scale with complete underlying driving processes (driving decisions, heating decisions, and so on).

One of the biggest challenges in visualizing the Vulcan data was obtaining interactive rates over multiple time steps. The time to generate the visualizations' geometry was limited by memory access time. The data set size often limited loading all the data into main memory, so we implemented an out-of-core visualization mechanism. Even loading only the data subset for each visualization still created a bottleneck per time step. The next step is to explore methods to facilitate interactivity.

References

1. K.R. Gurney et al., "Research Needs for Process-Driven, Finely Resolved Fossil Fuel Carbon Dioxide Emissions," *EOS Trans.*, vol. 88, no. 49, 2007; doi: 10.1029/2007EO490008.
2. K.R. Gurney et al., "Sensitivity of Atmospheric CO_2 Inversion to Seasonal and Interannual Variations in Fossil Fuel Emissions," *J. Geophysical Research*, vol. 110, no. D10, 2005, pp. 10,308–10,321.
3. A.S. Denning et al., *Science Implementation Strategy for the North American Carbon Program: Report of the NACP Implementation Strategy Group of the US Carbon Cycle Interagency Working Group*, US Carbon Cycle Science Program, 2005.
4. D. Crisp et al., "The Orbiting Carbon Observatory (OCO) Mission," *Advances in Space Research*, vol. 34, no. 4, 2004, pp. 700–709.
5. R.J. Andres et al., "A 1 × 1 Distribution of Carbon Dioxide Emissions from Fossil Fuel Consumption and Cement Manufacture, 1950–1990," *Global Biogeochemical Cycles*, vol. 10, no. 3, 1996, pp. 419–429.
6. A.L. Brenkert, "Carbon Dioxide Emission Estimates from Fossil-Fuel Burning, Hydraulic Cement Production, and Gas Flaring for 1995 on a One Degree Grid Cell Basis," 1998; http://cdiac.esd.ornl.gov/ndps/ndp058a.html.
7. D. Borland and R.M. Taylor, "Rainbow Color Map (Still) Considered Harmful," *IEEE Computer Graphics and Applications*, vol. 27, no. 2, 2007, pp. 14–17.

Nathan Andrysco *is a PhD student in Purdue University's Department of Computer Science. Contact him at nandrysc@purdue.edu.*

Kevin Robert Gurney *is an assistant professor in Purdue University's Department of Earth and Atmospheric Sciences and Department of Agronomy. Contact him at kgurney@purdue.edu.*

Bedřich Beneš *is an assistant professor in Purdue University's Department of Computer Graphics Technology. Contact him at bbenes@purdue.edu.*

Kathy Corbin *received her PhD from Colorado State University and is now serving as a postdoctoral researcher with the Commonwealth Scientific and Industrial Research Organization Marine and Atmospheric Research. Contact her at katherine.corbin@csiro.au.*