ATMOSPHERIC SCIENCE

Recording Earth's Vital Signs

Fifty years ago, continuous measurements of atmospheric carbon dioxide were begun at Mauna Loa, Hawaii.

Ralph F. Keeling

This year marks the 50th anniversary of the start of the Mauna Loa CO_2 record, the longest continuous record of CO_2 in the atmosphere. Initiated by my father, Charles D. Keeling of the Scripps Institution of Oceanography, the record provided the first compelling evidence that the concentration of CO_2 in the atmosphere was rising. It has become an icon of the human imprint on the planet and a continuing resource for the study of the changing global carbon cycle. The Mauna Loa story (*1*) provides a valuable lesson on the importance of continuous Earth observations in a time of accelerating global change.

At the outset, the decision to place the instrument at Mauna Loa was a gamble. Existing measurements suggested that atmospheric CO_2 concentrations varied widely depending on the place and time. Given this variability, could a meaningful record be recovered from an instrument parked in one location? Among the skeptics was Roger Revelle, then director of the Scripps Institution of Oceanography. Revelle would eventually become one of the record's strongest champions. Initially, however, he urged that priority be given to a one-time survey of CO_2 variability using ships and airplanes. Such a survey could be repeated a decade or so later to look for long-term changes.

My father was armed with evidence from his postdoctoral research that the CO_2 concentration in the remote atmosphere was a lot less variable than previously believed (*2*). He also had a strong ally in Harry Wexler of the U.S. Weather Bureau, who envisioned a central role for the newly established Mauna Loa Observatory (see the photo) in the major field program planned for the International Geophysical Year of 1957–1958.

The value of the Mauna Loa data soon was apparent (*3*). By the second year, a regular seasonal cycle was evident, reflecting the "breathing" of land plants in the Northern Hemisphere. Together with a more limited CO_2 data set from the South Pole, begun in 1957, the record documented a global rising trend attributable to the burning of fossil fuels worldwide (see the graph). In the 1960s and 1970s, the curve was seen by countless scientists, some of whom were drawn to study the science of global warming by the curve's ominous rise.

What if CO_2 had been measured only via repeated global surveys, as envisioned by Revelle? As the inset in the graph shows, a CO_2 record degraded to include only one point every decade or two loses its convincing message. Variations from survey to survey may be instrumental artifacts, or the apparent trend may be a random fluctuation. As a recent study of ocean currents in the North Atlantic has shown, resolving trends from repeated surveys can be perilously difficult (*4*).

The Scripps CO_2 program was shut down briefly in 1964 following congressionally mandated budget cuts. A more serious challenge loomed in the 1970s, when my father was asked to draw a line between the part of the CO_2 program that was basic research and the part that constituted "routine monitoring"; the routine activities would be transferred to a government agency. My father did not comply with the request.

In the 1970s, the Scripps CO_2 program expanded to an array of eight stations distributed globally. A large part of the effort was being expended not in routine data collection, but in the messier process of identifying and eliminating systematic errors. As the records grew, additional features emerged, such as a link between interannual CO_2 fluctuations and El Niño events (*5*) and changes in the amplitude of the seasonal cycle with time (*6*). Were
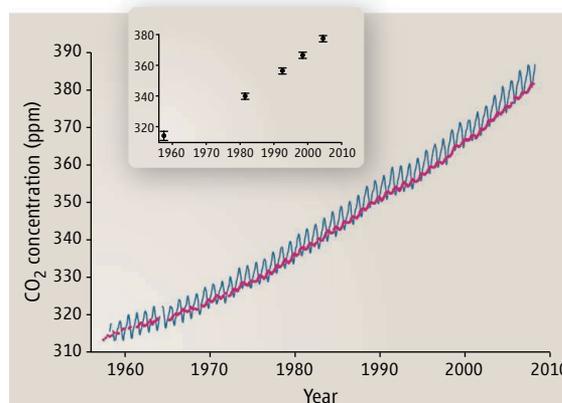
The value of continuous data. Within a few years, the continuous Mauna Loa (blue) and South Pole (red) records provided convincing evidence that CO_2 was rising. If CO_2 had been measured only as often as surveys of the North Atlantic overturning circulation (*4*), it would have taken decades to obtain convincing evidence (**inset**).

these features real or artifacts? There was no way to be sure without revisiting the fundamentals of the instrumentation, performing new calibrations, and reprocessing all of the records. This cycle was repeated many times as scientific interests evolved.

The distinction between research and routine monitoring may seem clear when applied to an activity like weather forecasting, but in the case of a program aimed at tracking long-term change, "research" and "operations" cannot be separated cleanly. Finding and correcting for the inevitable systematic biases is a job for scientists who understand the measurement technology, are passionate about data integrity, and are motivated to unravel how the Earth system operates.

The Mauna Loa experience also illustrates the critical need for redundancy. From the outset the Mauna Loa record was backed by the parallel record from the South Pole. In 1960, a second record was begun at Mauna Loa, based on flasks shipped back to Scripps for analysis. It is an inescapable fact that if you are trying to track changes over time, you only get one chance to measure each point. To prove you got it right, you must take measurements in multiple ways. And the challenge may come decades later. A recently
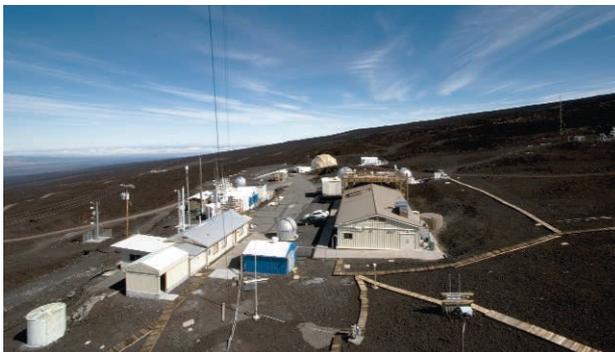
The Mauna Loa Observatory.

Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093, USA. E-mail: rkeeling@ucsd.edu

discovered slowing of the rate of CO_2 uptake by the Southern Ocean appears to hinge on questionable CO_2 measurements made at Ascension Island in the 1980s (7, 8). A duplicate record would have settled the issue.

The Scripps CO_2 program is now a component of a multinational collaboration aimed at tracking changes in greenhouse gases and related species, coordinated by the World Meteorological Organization. A major justification of this effort is the promise of quantifying the sources and sinks of greenhouse gases at Earth's surface. Soon, the compliance of international treaties to curb greenhouse gas emissions may be assessed using these capabilities.

If long-term observations are fundamental to understanding global change, why have they proved so hard to support? The costs of sustained measurements can be high, so prioritization is clearly an issue. The Scripps program has proved, however, that a long-term

observational program is not necessarily incompatible with the normal peer review system. The Scripps program continues to be funded—if perilously—one grant at a time. Even within agencies committed to long-term observations, such as the National Oceanographic and Atmospheric Administration, funding is tight and a hiatus may be only one political wind shift or economic downturn away. A diversity of funding sources supporting a heterogeneous mixture of overlapping programs is probably the best formula for long-term stability.

A continuing challenge to long-term Earth observations is the prejudice against science that is not directly aimed at hypothesis testing. At a time when the planet is being propelled by human action into another climate regime with incalculable social and environmental costs, we cannot afford such a rigid view of the scientific enterprise. The only way to figure out what

is happening to our planet is to measure it, and this means tracking changes decade after decade and poring over the records. A point of diminishing scientific returns has never been realized in what is now known as the "Keeling Curve," the Mauna Loa CO_2 record.

References
1. C. D. Keeling, *Annu. Rev. Energy Environ.* **23**, 25 (1998).
2. C. D. Keeling, in *Proceedings of the Conference on Recent Research in Climatology*, H. Craig, Ed. (Committee on Research in Water Resources and University of California, Scripps Institution of Oceanography, La Jolla, CA, 1957), pp. 43–49.
3. C. D. Keeling, *Tellus* **12**, 200 (1960).
4. H. L. Bryden, H. R. Longworth, S. A. Cunningham, *Nature* **438**, 655 (2005).
5. R. B. Bacastow, *Nature* **261**, 116 (1976).
6. R. B. Bacastow, C. D. Keeling, T. P. Whorf, *J. Geophys. Res.* **90**, 10529 (1985).
7. R. M. Law, R. J. Matear, R. J. Francey, *Science* **319**, 570a (2008).
8. C. Le Quéré *et al.*, *Science* **319**, 570c (2008).

BIOCHEMISTRY

A Postgenomic Visual Icon

A decade of experience in visualizing large-scale genotypic and phenotypic data as heat maps has illuminated the strengths and limitations of the approach.

John N. Weinstein

The "postgenomic era" in biology may be hard to define, and when it actually began is subject to debate. But its most characteristic feature is clearly the accumulation of massive amounts of genotypic and phenotypic data that must be organized, analyzed, visualized, and interpreted. That series of challenges has been central to recent bioinformatics. For visualization, by far the most popular graphical representation has been the "clustered heat map," which compacts large amounts of information into a small space to bring out coherent patterns in the data. Despite its popularity, however, are such maps optimal for visually integrating information to extract valuable insights and generate fresh hypotheses? That question can be addressed through understanding the strengths and limitations of heat map visualization.

Since their debut over 10 years ago (1) (see the figure), clustered heat maps have appeared in well over 4000 biological or biomedical publications. They have been used for two-dimensional display of patterns in all types of molecular data, including messenger RNA (mRNA) and microRNA expression, protein

expression, DNA copy number, DNA methylation, metabolite concentration, and drug activity (1–8). They have proved useful for microarray data (2) and have sometimes been engineered for "integromic" merging (1, 9, 10) of different types of molecular information. The figure, for example, combines data on mRNA expression, protein expression, mutations, cell cycle properties, stress responses, a yeast-based functional assay, and drug activity in cancer cells. Organisms analyzed have spanned the phylogenetic tree from the plant *Arabidopsis thaliana* to rainbow trout to suicidal crickets (11–13). Diseases analyzed have ranged from AIDS to cancer to bubonic plague (1, 14, 15).

In the case of gene expression data, the color assigned to a point in the heat map grid indicates how much of a particular RNA or protein is expressed in a given sample. The gene expression level is generally indicated by red for high expression and either green or blue for low expression. Coherent patterns (patches) of color are generated by hierarchical clustering on both horizontal and vertical axes to bring like together with like. Cluster relationships are indicated by tree-like structures adjacent to the heat map, and the patches of color may indicate functional relationships among genes and samples. Occasionally, a

source of order other than clustering (for example, time in a series of measurements) is used on one or both axes. Without some basis for functional ordering on both axes, however, there would be no coherent patterns of color.

Seductive though it may be, the clustered heat map has its limitations and potential for misinterpretation or misuse. Most prominently among the limitations, it provides only first-order insight into the data; complex patterns of nonlinear relationship among only a few of the samples are unlikely to show up. A computer-intensive variant based on "biclustering" has been developed to reveal such relationships (16). A second problem is that, in hierarchical clustering, each bifurcation of the cluster tree can be "swung" in either direction at each fork in the tree, so some objective (but, to a degree, arbitrary) rule must be invoked to decide which way each branch will, in fact, swing. There is also the temptation to select a small subset of the variables (for example, genes in a microarray study), and represent them in a clustered heat map. That is common (and appropriate) practice in the discovery of new biomarkers and gene expression signatures for discriminating subtypes of a disease such as cancer (17). However, if one picks a signature consisting of only a few dozen genes out of a set of more than 10,000, then even randomized

Department of Bioinformatics and Computational Biology, M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. E-mail: jweinste@mdanderson.org