method work effectively. They do this by making good use of recently developed machine learning algorithms for a special class of neural networks (*5*, *6*).

Hinton and Salakhutdinov's approach uses so-called autoencoder networks—neural networks that learn a compact description of data, as shown in the middle panel of the figure. This is a neural network that attempts to learn to map the three-dimensional data from the spiral down to one dimension, and then back out to three dimensions. The network is trained to reproduce its input on its output—an identity mapping—by the standard backpropagation of error method (*7*, *8*). Although backpropagation is a supervised learning method, by using the input as the teacher, this method becomes unsupervised (or self-supervised). Unfortunately, this network will fail miserably at this task, in much the same way that standard methods such as principal components analysis will fail. This is because even though there is a weighted sum of the inputs (a linear mapping) to a representation of x—the location along the spiral—there is no (semi-)linear function (*9*) of x that can decode this back to $\sin(x)$ or $\cos(x)$. That is, the network is incapable of even representing the transformation, much less learning it. The best such a network can do is to learn the average of the points, a line down the middle of the spiral. However, if another nonlinear layer is added between the output and the central hidden layer (see the figure, right panel), then the network is powerful enough, and can learn to encode the points as one dimension (easy) but also can learn to decode that one-dimensional representation back out to the three dimensions of the spiral (hard). Finding a set of connection strengths (weights) that will carry out this learning problem by means of backpropagation has proven to be unreliable in practice (*10*). If one could initialize the weights so that they are near a solution, it is easy to fine-tune them with standard methods, as Hinton and Salakhutdinov show.

The authors use recent advances in training a specific kind of network, called a restricted Boltzmann machine or Harmony network (*5*, *6*), to learn a good initial mapping recursively. First, their system learns an invertible mapping from the data to a layer of binary features. This initial mapping may actually increase the dimensionality of the data, which is necessary for problems like the spiral. Then, it learns a mapping from those features to another layer of features. This is repeated as many times as desired to initialize an extremely deep autoencoder. The resulting deep network is then used as the initialization of a standard neural network, which then tunes the weights to perform much better.

This makes it practical to use much deeper networks than were previously possible, thus allowing more complex nonlinear codes to be learned. Although there is an engineering flavor to much of the paper, this is the first practical method that results in a completely invertible mapping, so that new data may be projected into this very low dimensional space. The hope is that these lower dimensional representations will be useful for important tasks such as pattern recognition, transformation, or visualization. Hinton and Salakhutdinov have already demonstrated some excellent results in widely varying domains. This is exciting work with many potential applications in domains of current interest such as biology, neuroscience, and the study of the Web.

Recent advances in machine learning have caused some to consider neural networks obsolete, even dead. This work suggests that such announcements are premature.

References and Notes

1. G. E. Hinton, R. R. Salakhutdinov, *Science* **313**, 504 (2006).
2. S. T. Roweis, L. K. Saul, *Science* **290**, 2323 (2000).
3. J. A. Tenenbaum, V. J. de Silva, J. C. Langford, *Science* **290**, 2319 (2000).
4. One can learn a mapping to the manifold (and back), but this is done independently of the original structure-finding method, which does not provide this mapping.
5 G. E. Hinton, *Neural Comput.* **14**, 1771 (2002).
6. P. Smolensky, in *Parallel Distributed Processing*, vol. 1, *Foundations*, D. E. Rumelhart, J. L. McClelland, PDP Research Group, Eds. (MIT Press, Cambridge, MA, 1986), pp. 194–281.
7. D. E. Rumelhart, G. E. Hinton, R. J. Williams, *Nature* **323**, 533 (1986).
8. G. W. Cottrell, P. W. Munro, D. Zipser, in *Models of Cognition: A Review of Cognitive Science*, N. E. Sharkey, Ed. (Ablex, Norwood, NJ, 1989), vol. 1, pp. 208–240.
9. A so-called semilinear function is one that takes as input a weighted sum of other variables, and applies a monotonic transformation to it. The standard sigmoid function used in neural networks is an example.
10. D. DeMers, G. W. Cottrell, in *Advances in Neural Information Processing Systems*, S. J. Hanson, J. D. Cowan, C. L. Giles, Eds. (Morgan Kaufmann, San Mateo, CA, 1993), vol. 5, pp. 580–587.

10.1126/science.1129813

ATMOSPHERE

What Drives the Ice Age Cycle?

Between 3 and 1 million years ago, ice ages followed a 41,000-year cycle. Two studies provide new explanations for this periodicity.

Didier Paillard

The exposure of Earth's surface to the Sun's rays (or insolation) varies on time scales of thousands of years as a result of regular changes in Earth's orbit around the Sun (eccentricity), in the tilt of Earth's axis (obliquity), and in the direction of Earth's axis of rotation (precession). According to the Milankovitch theory, these insolation changes drive the glacial cycles that have dominated Earth's climate for the past 3 million years.

For example, between 3 and 1 million years before present (late Pliocene to early Pleistocene, hereafter LP-EP), the glacial oscillations followed a 41,000-year cycle. These oscillations correspond to insolation changes driven by obliquity changes. But during this time, precession-driven changes in insolation on a 23,000-year cycle were much stronger than the obliquity-driven changes. Why is the glacial record for the LP-EP dominated by obliquity, rather than by the stronger precessional forcing? How should the Milankovitch theory be adapted to account for this "41,000-year paradox"?

Two different solutions are presented in this issue. The first involves a rethinking of how the insolation forcing should be defined (*1*), whereas the second suggests that the Antarctic ice sheet may play an important role (*2*).The two papers question some basic principles that are often accepted without debate.

On page 508, Huybers (*1*) argues that the summer insolation traditionally used in ice age models may not be the best parameter. Because ice mass balance depends on whether the temperature is above or below the freezing point, a physically more relevant parameter should be the insolation integrated over a given threshold that allows for ice melting. This new parameter more closely follows a 41,000-year periodicity, thus providing a possible explanation for the LP-EP record.

On page 492, Raymo *et al*. (*2*) question another pillar of ice age research by suggesting that the East Antarctic ice sheet could have contributed substantially to sea-level changes during the LP-EP. The East Antarctic ice sheet is land-based and should therefore be sensitive mostly to insolation forcing, whereas the West Antarctic ice sheet is marine-based and thus influenced largely by sea-level changes. Because the obliquity forcing is symmetrical with respect to the hemispheres, whereas the preces-

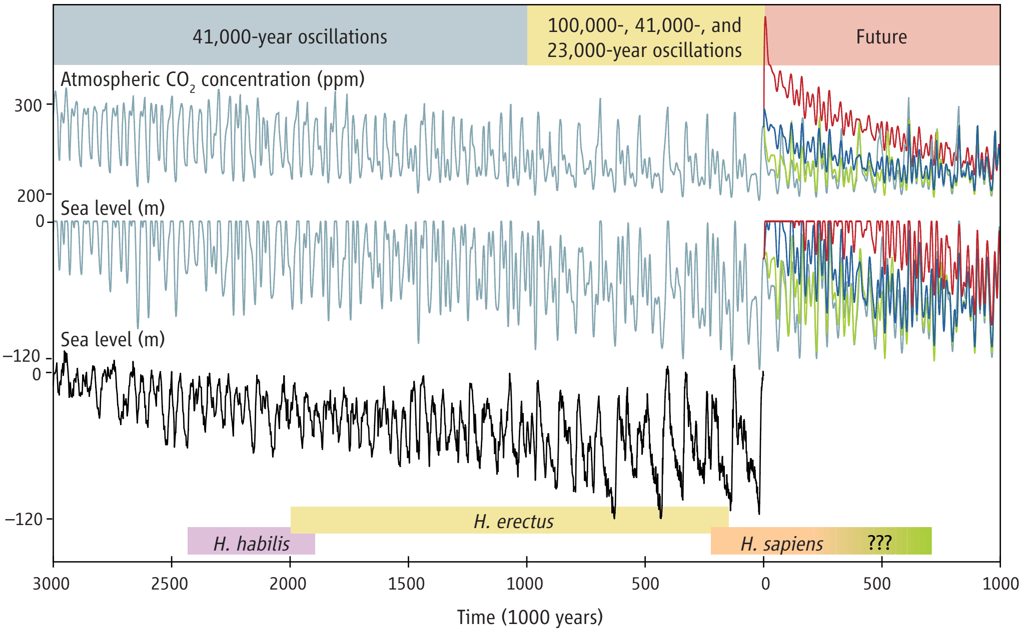
---

The author is at the Laboratoire des Sciences du Climat et de l'Environnement, Institut Pierre Simon Laplace, CEA-CNRS-UVSQ, 91191 Gif-sur-Yvette, France. E-mail: didier.paillard@cea.fr

Published by AAAS

sional forcing is antisymmetrical, the contributions of the northern and southern ice sheets to the global ice volume record will add up for the 41,000-year cycle, but cancel each other out for the 23,000-year cycle, thus explaining the 41,000-year paradox.

Both hypotheses could be part of the solution. Huybers's idea is based on a sound and simple physical premise and is certainly valid to some extent. The hypothesis of Raymo *et al.* provides a scenario for an increasing contribution of the 23,000-year cycles under a colder climate, through a transition from a land-based to a marine-based East Antarctic ice sheet around 1 million years ago. Indeed, though not dominant, the precessional cycles are present in the climate record of the past 1 million years (the late Pleistocene). Still, neither hypothesis can account for the beginning of Northern Hemisphere glaciations around 3 million years ago. Furthermore, during the past 1 million years, glacial-interglacial oscillations have largely been dominated by a 100,000-year periodicity, yet there is no notable associated 100,000-year insolation forcing. There is currently no consensus on what drives these late Pleistocene 100,000-year cycles.



Past and future climate. Simulated cycles of atmospheric CO$_2$ concentrations (**top**) and sea level (**middle**) from 3 million years before present to 1 million years in the future (*5*). The model accounts for the interaction between ice volume and atmospheric CO$_2$ concentrations. The amplitude of future climatic cycles may share similarities with those in the late Pliocene (about 3 million years ago), depending on the total amount of CO$_2$ released into the atmosphere through human activities (*8*). Gray: without anthropogenic CO$_2$ emissions; green: 450 gigatons of carbon (GtC), assuming that emissions stop today; blue: 1500 GtC, an optimistic emissions scenario; red: 5000 GtC, a pessimistic emissions scenario, assuming that the entire estimated reservoir of fossil fuels on Earth is burnt. (**Bottom**) Isotopic record of past ice volume, showing 41,000-year cycles between 3 and 1 million years ago and larger 100,000-year cycles since 1 million years ago (*9*).

The theories of Huybers and Raymo *et al.* can be traced back to the 19th century. In 1842, Adhémar proposed that the ice ages were driven by precessional changes (obliquity and eccentricity changes were unknown at this time). Because precessional changes are antisymmetrical with respect to the hemispheres, he argued that Antarctica is glaciated today, whereas some time ago, the northern hemisphere was covered by ice, thus explaining the geologic field data (*3*). This alternation between the hemispheres is somewhat like in (*2*). His theory was dismissed at the time by Lyell and by Alexander von Humboldt (*3*), because the amount of energy received on Earth does not depend on precession: more intense (colder) winters were also shorter, with the energy budget at the top of the atmosphere being unchanged because precession modulates not only the intensity but also the duration of seasons. Precession should thus not affect climate, somewhat like in (*1*).

Since the 19th century, two families of ice age theories have been put forward: insolation-based theories proposed by Adhémar, Croll, and Milankovitch, and atmospheric CO$_2$ ones proposed by Tyndall, Arrhenius, and Chamberlin (*3*). The latter theories suggested that glaciations were associated with lower CO$_2$ levels. This is now confirmed by the large oscillations in atmospheric CO$_2$ measured in Antarctic ice cores over the past 650,000 years (*4*). It is certainly difficult to explain the ice ages of the past 1 million years purely on the basis of insolation changes. In the late Pleistocene, both insolation changes and atmospheric CO$_2$ concentrations must have played a critical role in the dynamics of glaciations, although a final synthesis still eludes us.

The big challenge is to build an ice age theory that can account not only for ice sheet and atmospheric CO$_2$ changes, but also for the start of glaciations about 3 million years ago and for the transition from 41,000-year cycles to much larger 100,000-year oscillations around 1 million years ago. The atmospheric CO$_2$ concentration was probably very important over the past 1 million years, but was this also the case during the LP-EP? Alternatively, if one can build a purely insolation-based theory between 3 and 1 million years ago, as suggested by Huybers and Raymo *et al.*, why is this not the case anymore in the past 1 million years?

A tentative scenario, based on a bistable ocean system (*5*), is shown in the figure, where the 41,000-year paradox and the 100,000-year problem have a common answer in an oceanic switch that can store or release carbon depending on ice-sheet size and insolation forcing, using empirical relationships. This conceptual model can be extrapolated to a future with and without anthropogenic CO$_2$ emissions. The results are comparable to those of more sophisticated models (*6*), providing a framework for understanding the likely climatic future of our planet in the context of the climate of the past 3 million years.

The mid-Pliocene, about 3.3 to 3.0 million years ago, has been cited as a possible analog for our future warmer Earth (*7*). This and the subsequent LP-EP time period are interesting not only in terms of their climate, but also because during this period, *Homo habilis* first appeared on the scene. Furthermore, they are currently our best guide to what climate and ice sheets may look like for *Homo sapiens* to come. The reports by Huybers and by Raymo *et al.* bring us a step closer to understanding the dynamics of these past climates.

References and Notes

1. P. Huybers, *Science* **313**, 508 (2006).
2. M. E. Raymo, L. E. Lisiecki, K. H. Nisancioglu, *Science* **313**, 492 (2006).
3. E. Bard, *C. R. Geosci.* **336**, 603 (2004).
4. U. Siegenthaler *et al.*, *Science* **310**, 1313 (2005).
5. D. Paillard, F. Parrenin, *Earth Planet. Sci. Lett.* **227**, 263 (2004).
6. D. Archer, A. Ganopolski, *Geochem. Geophys. Geosyst.* **6**, Q05003 (2005).
7. H. Dowsett *et al.*, *Global Planet. Change* **9**, 169 (1994).
8. To build this figure, the model in (*5*) was extrapolated using a decay *e*-folding time of 400,000 years for the removal by silicate weathering of a remaining 8% long-lived part of total anthropogenic carbon, following (*10*).
9. L. E. Lisiecki, M. E. Raymo, *Paleoceanography* **20**, PA1003 (2005).
10. D. Archer, *J. Geophys. Res.* **110**, C09S05 (2005).

10.1126/science.1131297