Epsilon Machine Computation in Paleoclimate Proxy Data: Implications for Testing Climate Models

Michael R. Gipp*
Marine Mining Corp., 5829 Fieldon Rd., Mississauga, ON L5M 5K3
drgipp@gmail.com

Summary

The problem of describing the behaviour of global climate and producing a convincing model of the same is a problem of computational mechanics. One approach to describing complex systems is Crutchfield's (1994) epsilon machine concept, which is an iterative approach to describing a data set, such as a paleoclimatic proxy or a model output, in terms of automata of successively greater complexity. Two paleoclimatic data sets (deep marine $\delta^{18}O$, and the magnetic susceptibility of loess) are studied: first by reconstructing their phase space portraits in two dimensions; then by contouring the probability density of the phase space portraits over successive windows in time. Areas of consistently high probability across several time windows represent areas of stability. The epsilon machine is defined by the sequence in which the areas of stability are visited during the forward time evolution of the phase space portrait. The same approach, applied to climate models, shows that most do not reach the level of complexity of the paleoclimate record, with but few exceptions: 1)the model of Paillard (2001), which shows some complexity, but does so by explicitly defining in the model; and 2) the model of Saltzman and Verbitsky (1994), in which complex behaviour appears to be an emergent property of the model, although the level of complexity falls short of that of the paleoclimate proxies.

Introduction

Climate modeling is critically important for forecasting impacts of climate change on human society and the natural world. One of the most significant problems with climate modelling is the difficulty of testing such models. Nonlinear terms within the model will cause errors in the initial conditions to progressively increase until the expected correlation between a model output and a real record falls to zero. Testing must instead be carried out by comparing invariant properties of the climate model outputs with similar properties inferred from paleoclimatic proxies. Characterizing the behavior of the Quaternary climate system and modeling the same are thus two sides of the same problem.

In this paper the author applies the epsilon machine concept to characterize the behaviour of the climate system. The data studied in paper includes the deep-sea $\delta^{18}O$ record of ODP 677 (Shackleton *et al.,* 1990) and the record of magnetic susceptibility of loess and paleosols at Luochuan, central China (Kukla *et al*., 1990). Both of these records are long (>1.8 Ma) and cover the most significant portion of the Quaternary.

Theory and Discussion

The climate system is normally regarded as a damped driven deterministic system (e.g., Saltzman, 2002). Such a deterministic system can be defined by a phase space with coordinates $x_1, x_2, x_3, \ldots x_n$, and the functions $x_1(t), x_2(t), x_3(t), \ldots, x_n(t)$ describe its time evolution. The number of output functions (*n*) is called the embedding dimension (Sauer *et al*., 1991). The evolution of the system is marked by the trajectory traced out by sequential plots of individual states with coordinates defined by the values of the *n* functions at each observed time. The coordinates may each represent an independent measurement, successively higher time derivatives of $x_1(t)$, or lagged observations (e.g., $x_1(t)$, $x_1(t+\Delta t)$, $x_1(t+2\Delta t), \ldots$).

Ergodic theory suggests that dynamic information about the entire system is contained in each time series output from the system (Abarbanel, 1996). Therefore, a phase space portrait reflecting the dynamics of the entire system may be reconstructed from a single time series. The phase space is reconstructed by the time delay method (Packard *et al*., 1980), in which two-dimensional phase space portraits were created by plotting each time series against a lagged observation from the same series.

Using the method of Gipp (2001; in press) to create probability density plots of in reconstructed two-dimensional phase space portraits of both records, we identify six regions of Lyapunov stability, which, in the case of the ice volume proxy, represent stable ice configurations that recur episodically throughout the Quaternary (figure 1).

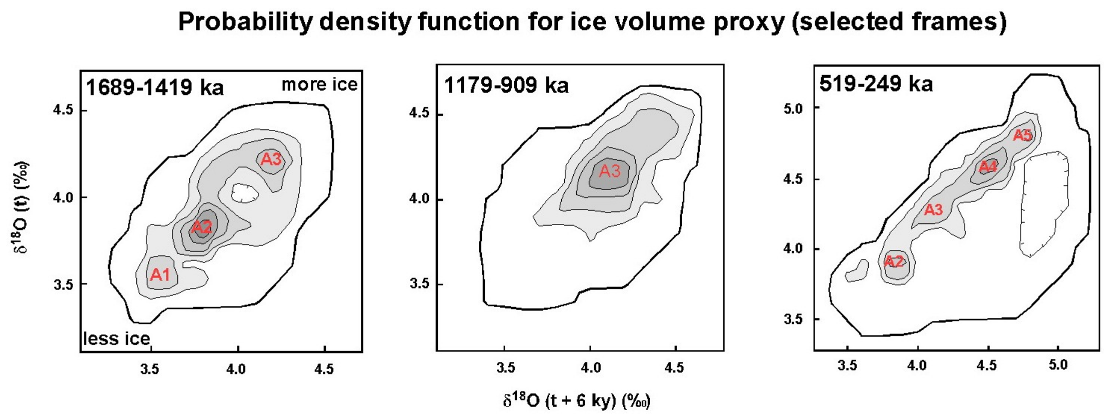

Figure 1: Three selected frames from the animation of the probability density of the ice volume reconstructed phase space portrait. Note the change in scale from left to right. In the early Quaternary (at left), there were three stable configurations for ice: an interglacial state (A1) representing considerably less ice than at present, a mid-glacial state (A2) representing an ice volume comparable to that of today, and a glacial maximum (A3) representing slightly more ice than at present. A typical state sequence in the early Quaternary would run A1-A2-A3-A2-A1- . . .For a brief time at the beginning of the mid-Pleistocene transition (centre), there was only one stable ice configuration (A3). Glacial cycles were short, consisting of slow growth to a point of instability, followed by rapid deglaciation to another point of instability, then rapid glaciation back to A3. By the late Quaternary (at right), there are four stable ice states: interglacial (A2, roughly the ice volume at present); two mid-glacial states (A3 and A4) and a maximum glacial state (A5, similar to the last glacial maximum). A typical state sequence is A2-A3-A4-A5-A2- . . . resulting in the characteristic sawtooth pattern of late Quaternary glaciations. After Gipp (in press).

The paleoclimate proxy for paleomonsoon strength similarly shows six regions of Lyapunov stability (labeled M1 to M6 in figure 2). Each region represents a relatively stable strength for the Himalayan monsoon on a millennial timescale.

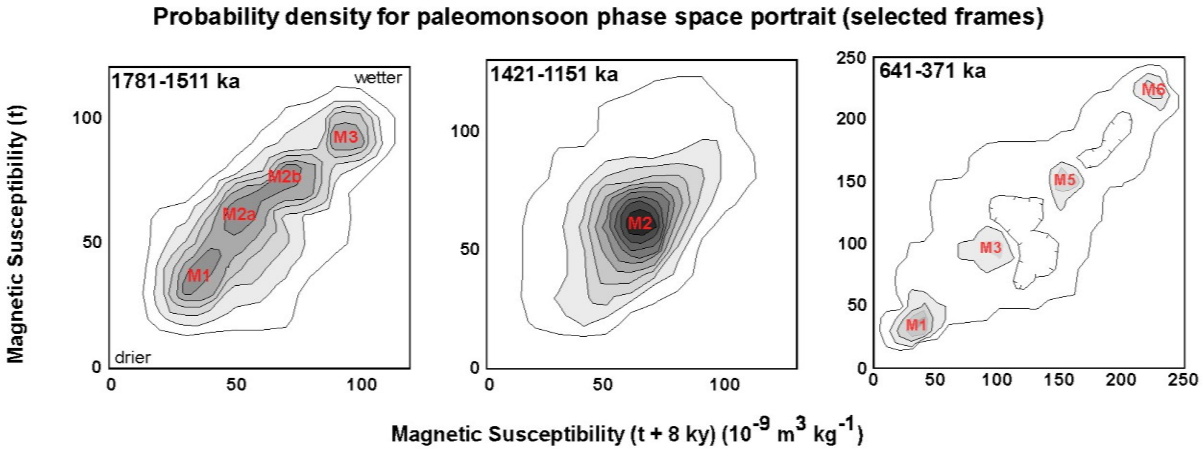

Figure 2. Three selected frames from the animation of the probability density plot of the Himalayan paleomonsoon phase space portrait. Note the scale change from left to right. In the early Quaternary (left) there were three stable monsoonal strengths: M1, which is comparable to the Himalayan monsoonal strength during the last glacial maximum; M2 (here bifurcated into M2a and M2b), comparable to recent strength; and M3, last seen at about 75 ka. The splitting of M2 into M2a and M2b happens twice during the Quaternary, first at about 1800 ka and again at about 1220 ka. Near the beginning of the Mid-Pleistocene transition (middle) there is one stable state—a feature that persists for roughly 200 ky. Near the end of the Mid-Pleistocene transition (right) there is a rapid expansion of the function into previously inaccessible areas of phase space.

The climate system has been widely recognized as being a complex system; however characterizing that complexity may be a subjective exercise. The epsilon machine concept of Crutchfield (1994) is here proposed as a means of analyzing the complexity of the climate system in a less subjective manner. The epsilon machine concept was presented as an iterative method where complexity is defined within a data stream at progressively higher levels. The data are studied for the presence of similar “causal states”, in which the data evolve in a similar fashion through time. As all similar causal states are identified, the sequence of causal states defines a low-ordered automaton (a second-order epsilon machine), which is a statement of the complexity of the system.

Rigorous definition of causal states from paleoclimate data is difficult, because the record is neither long enough nor sufficiently repetitive to define any conclusively similar states. However, the sequence of zones of stability (A1-A6 for the ice volume phase space portrait; M1-M6 for the paleomonsoon strength phase space portrait) can be used to define a second-order epsilon machine.

Furthermore, the structure of these epsilon machines changes abruptly (or bifurcates) numerous times over the past 1.8 million years, allowing us to define higher level “causal states”, the sequences of which define a third-order epsilon machine. Whereas the lowest order of complexity is driven by the response of coexisting negative and positive feedbacks within a system undergoing astronomical forcing; higher ordered complexity arises from the time variation of boundary conditions (including CO_2) which, over the period represented by the data seta, have a linear or sublinear trend superimposed on quasiperiodic variation.

Larger scales of cyclicity in the boundary conditions may exist due to episodes of continental collisions and cycles of supercontinent formation, which may create higher levels of complexity. In this way, fourth- and fifth-order epsilon machines could potentially be defined within paleoclimate records of sufficient length.

The output of climate models can be characterized using the epsilon machine approach and the results compared with the paleoclimate record to see whether the complexity in the geologic record is present in the model. Investigating the models of Paillard (2001) and process models (Saltzman and Verbitsky, 1994; Saltzman, 2002) yields interesting results. The Paillard model was explicitly designed as a second-order epsilon machine, and consequently does not yield any higher order complexity. The Saltzman model is designed on the basis of assumed physical properties, and its output resembles a second order epsilon machine; and it is short enough that it does not cover any portion of the climate record where bifurcations and hence third order epsilon machines are apparent. But as this model was not specifically designed to operate as a second- or higher-order epsilon machine; the fact that it does makes it very interesting.

References

Abarbanel, H. D. I., 1996. Analysis of Observed Chaotic Data. Springer-Verlag, New York.

Crutchfield, J. P., 1994. The calculi of emergence: Computation, dynamics, and induction. *Physica* **D 75**: 11-54.

Gipp, M. R., 2001. Interpretation of climate dynamics from phase space portraits: Is the climate system strange or just different? *Paleoceanography*, **16**, 335-351.

Gipp, M. R., in press. Imaging system change in reconstructed phase space: Evidence for multiple bifurcations in Quaternary paleoclimate. G3.

Kukla, G., Z. S. An, J. L. Melice, J. Gavin, and J. L. Xiao, 1990. Magnetic susceptibility record of Chinese loess. *Trans. R. Soc. Edinburgh Earth Sci.,* **81:** 263-288.

Packard, N. H., J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, 1980. Geometry from a time series, *Phys. Rev. Lett., 45*, 712-716.

Paillard, D., 2001. Glacial cycles: Toward a new paradigm. *Reviews of Geophysics*, **3**: 325-346.

Saltzman, B., 2002. Dynamical Paleoclimatology: Generalized Theory of Global Climate Change. Academic Press, San Diego, 320 pp.

Saltzman, B., and Verbitsky, M., 1994. Late Pleistocene climatic trajectory in the phase space of global ice, ocean state, and CO_2: Observations and theory. *Paleoceanography*, **9:** 767–779.

Shackleton, N. J., A. Berger, and W. R. Peltier, 1990. An alternative astronomical calibration of the Lower Pleistocene timescale based on ODP site 677, *Trans. R. Soc. Edinburgh, Earth Sci.,* **81**: 251-261.