

QC  
807.5  
.U6  
W6  
no.297  
c.2

NOAA Technical Memorandum OAR ETL-297



---

## REPLACING MODEL PARAMETERIZATION WITH EPSILON MACHINES

J. Palmer

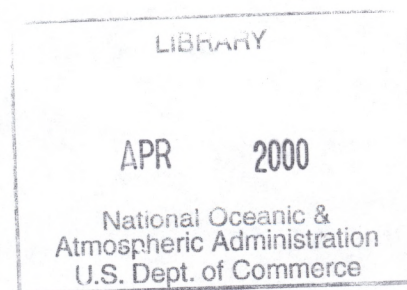
Environmental Technology Laboratory  
Boulder, Colorado  
October 1999

NOAA Technical Memorandum OAR ETL-297

## REPLACING MODEL PARAMETERIZATION WITH EPSILON MACHINES

*A Proposed Long-Term Project at the Environmental Technology Laboratory*

Jay Palmer



QC  
807.5  
U6  
W6  
no. 297  
C.2

Environmental Technology Laboratory  
Boulder, Colorado  
October 1999



**UNITED STATES  
DEPARTMENT OF COMMERCE**

**William M. Daley  
Secretary**

**NATIONAL OCEANIC AND  
ATMOSPHERIC ADMINISTRATION**

**D. JAMES BAKER  
Under Secretary for Oceans  
and Atmosphere/Administrator**

**Oceanic and Atmospheric  
Research Laboratories**

**David L. Evans  
Director**

## NOTICE

Mention of a commercial company or product does not constitute an endorsement by the NOAA Oceanic and Atmospheric Research Laboratories. Use of information from this publication concerning proprietary products or the test of such products for publicity or advertising purposes is not authorized.

---

For sale by the National Technical Information Service, 5285 Port Royal Road  
Springfield, VA 22061

# Replacing Model Parameterizations with Epsilon Machines

*A Proposed Long-Term Project at the Environmental Technology Laboratory*

Jay Palmer

October, 1999

## Contents

1.	Objective	1
2.	Background	1
3.	Epsilon machines as sub-grid models	2
4.	Performance comparisons	5
5.	Candidate sub-grid processes	5
6.	Summary	7
	References	7

## **1. Objective**

The objective of this research is to optimize the performance of numerical weather and climate models given fixed observational data on sub-grid processes and fixed computational memory for implementing the sub-grid model. This is accomplished by replacing physical parameterizations in the models with a type of finite state machine called an epsilon machine.

## **2. Background**

The importance of numerical climate modeling today demands that we strive to make optimal use of the computational resources applied to this task. A basic computational resource is the computer memory available during the computations. Currently, the atmosphere is modeled with a patchwork of deterministic and statistical modeling frameworks, and the required computer memory is simply allotted to meet the requirements of this modeling framework. A more scientific approach to the problem would be to let the observational data on the climate system tell us what memory is required for a given observational window on the data. Unfortunately, on scales smaller than the grid scale, the observational data is often immediately filtered into statistical means and moments in order to fit a favored "physical parameterization" and the memory content of the data is lost.

Models of sub-grid processes are required in all atmospheric and ocean computer models, e.g. global climate models, weather prediction models, mesoscale models, and even large-eddy simulation models. There are two components to all of these models. One component is termed the "dynamic" component, the other is termed the sub-grid or physical parameterization component. The dynamic component of the models is based on the Navier-Stokes equations of fluid dynamics. This portion of the model is supported by centuries of data extending back to the time of Newton, and it is extremely unlikely that new data can be used to improve upon it. On the other hand, the sub-grid parameterization portion of the models is ripe for improvement through the use of observational data, particularly if the data is used in new modeling frameworks capable of capturing "emergent" behavior resulting from nonlinear couplings between processes governed by the dynamic and sub-grid portions of the model.

### 3. Epsilon Machines as Sub-Grid Models

The *epsilon machine* is a symbolic dynamic modeling framework for locating the causal patterns in data originally represented as a string.<sup>1</sup> I propose the use of epsilon machines to model sub-grid processes because, for synthetic complex systems at least, the epsilon machine has been proven to be an optimal model. By optimal, I mean that the model yields the most accurate predictions possible with the minimum computational memory for a given observational window onto the system.

The procedure for constructing an epsilon machine from the data is termed “machine reconstructing” and is well documented in the literature.<sup>2</sup> An epsilon machine is a generalized Markov model represented by a particular type of finite state machine known as a deterministic, stochastic automaton. It can be classified as a Markov model because the output from the model depends only on its current state. It takes the form of a stochastic automaton because the transitions between the states are assigned transition probabilities. It is deterministic because a transition from a given state on a given symbol can be to only a single state. This last property is what allows the epsilon machine to serve as a deterministic algorithm without the need to invoke an ensemble average of machine runs.

Because the transitions between epsilon machine states are labeled with probabilities, the machine operation is both deterministic and probabilistic. The machine states themselves are certain invariant sub-trees found within the parse tree for the original data string. Because these sub-trees are, in general, branched structures, the machine states also have both deterministic and probabilistic structure. This dual structure in both the states and the state transitions, and the fact that this structure is *discovered* in the data rather than imprinted onto it, is what makes the epsilon machine modeling framework a more powerful and less biased modeling framework for sub-grid processes than the physical parameterization framework.

Finally, the computational memory of the epsilon machine can be computed as the entropy of the stationary states of the machine, and is known as the “statistical complexity” of the model.<sup>1,2</sup> This allows comparisons to be made with other sub-grid modeling frameworks under the constraint of each having the same memory allocation.

## Building the Epsilon Machine

The basic epsilon machine construction process is a “reconstruction” of data, first from a string, then to a parse tree, and finally to the epsilon machine.<sup>2</sup> Thus, the experimental data on the sub-grid process must eventually be put in the form of a string of symbols. The epsilon machine essentially identifies causal patterns in the string, i.e. sub-sequences in the string that are then put in correspondence with a machine state.

In the general, “multivariate” case, there will be several diagnostic variables for the sub-grid process, as well as several associated prognostic variables from the dynamic portion of the climate model. In the most direct epsilon machine approach to sub-grid modeling, a set of simultaneously *measured* values for all of these variables is first represented as a single string. If there is an underlying nonlinear dynamic system governing the sub-grid process that involves these variables, then a theorem of Kolmogorov implies that the most efficient partition for the data for describing this underlying dynamical system is the maximum entropy partition.<sup>3</sup> Finding this partition generally requires a global search process, and is the first step in building the epsilon machine model for the sub-grid process.<sup>4</sup> For  $n$ -dimensional data, the maximum entropy partition is found (approximately) by coarse graining the data with respect to a trial hyperplane partition of dimension  $n-1$ . In other words, the data is binned into a finite set of symbols, each of which corresponds to the position of the data point relative to the partition. The simplest coarse graining is to assign a 0 or 1 to the data point depending on whether a particular data point is on either side of the partition.<sup>4</sup>

This partitioning results in the representation of the sub-grid process data as a single time series, in the simplest case a string of 0s and 1s, that encodes the time series variability of the sub-grid process. Once this string has been produced, an epsilon machine that models the temporal evolution of the data can be constructed as described in the literature.<sup>2</sup> The causal states found for the machine will consist of patterns of partitioned data that will not be obvious from the form of the maximum entropy partition alone. “Dangling states” that result from nonstationarity of the data can be identified with the start state of the machine as the “state of total ignorance”.<sup>4</sup> This is a less biased measure of nonstationarity than existing measures that refer to statistical means and moments.

## Using the Epsilon Machine

The constructed epsilon machine can now be used as a sub-grid model in the following way. Say the dynamical portion of the larger model has reached a point where it requires a value for a sub-grid process variable. Consider, for example, that the sub-grid variable is surface stress. The dynamical portion of the large model has computed, among other variables, a value for the horizontal wind speed at a grid point. The standard sub-grid parameterization for surface stress simply multiplies this wind speed with a drag coefficient to obtain the needed sub-grid stress value.

In order for the epsilon machine model to provide a stress value, the current state of the sub-grid, wind-stress system, as described by the machine, must be known. If the epsilon machine model is being interrogated for the first time, then the current machine state is set to the start state. Next, the machine makes a transition from the start state in accordance with the transition probabilities specified for the machine using a weighted random number generator. The epsilon machine is left in a new state and has emitted a specified symbol in accordance with the transition matrix defined for the machine. This symbol encodes a sector of wind and stress data relative to the maximum entropy partition. The current value of the wind-speed from the dynamic portion of the model is then used to define a subset of stress values within this sector by referring to the original wind-stress data set. A median value from this subset of stress values is then given to the dynamic model.

If there are no data points in the identified sector at the current wind-speed value, then the epsilon machine is returned to the start state (state of total ignorance.) This action is an unbiased representation of the fact that the epsilon machine model of the wind-stress system and the overall model have lost synchronism. Otherwise, the machine simply makes the next transition in synchronism with the time step taken by the dynamic portion of the model. The implementation of this cooperative operation of the sub-grid model with the dynamical model is not possible with the standard sub-grid parameterization schemes.

#### 4. Performance Comparisons

Sub-grid models based on epsilon machine and on parameterizations have fundamentally different structures and philosophies. For a meaningful comparison to be made of these two methods, it is essential that we define a *conditional* comparison. As argued in the introduction, considering that sub-grid models consume most of the computer memory in climate and weather models, the most appropriate condition for comparison is that of *equal computational memory*. As mentioned above, the computational memory of an epsilon machine (statistical complexity) is well defined and it is easy to compute once an epsilon machine has been constructed. Most of the sub-grid parameterizations also have obvious memory allocation, primarily the memory needed to store the floating point variables in the algorithm. For the performance comparisons, the memory allocated to the parameterization algorithm is simply set equal to the statistical complexity of the epsilon machine plus the memory allocated to the reverse coarse graining procedure described above.

In making the comparisons of the two types of sub-grid models, their performance measure will be the performance of the model as a whole which, in turn, is measured by comparisons of the model output with observations. An ideal modeling platform for conducting this comparison is a platform designed expressly for the purpose of evaluating sub-grid models, known as single-column models. These models time-evolve only a single vertical column from the grid of columns that normally constitutes a full climate or weather model. The NCAR single column model known as the NCAR SCCM is a convenient well documented choice available on the Web.<sup>5</sup>

#### 5. Candidate Sub-Grid Process

This section lists a selection of three sub-grid processes currently modeled in the NCAR climate models<sup>6</sup> that are well suited for epsilon machine modeling using data currently archived at the Environmental Technology Laboratory (ETL). Epsilon machines constructed with these data sets are likely to lead to improved model performance because the data include temporal variabilities that occur over time scales comparable to the time interval that the dynamical portion of the models can resolve.

## **Clouds**

Currently cloud fraction and the associated optical properties are parameterized by a scheme that depends on instantaneous values of relative humidity, vertical velocity, atmospheric stability, and parameterized moist convection mass flux.<sup>6</sup> Replacing this scheme with one that depends on dynamical patterns of these quantities can be accomplished with an epsilon machine model constructed from data obtained at projects where ETL cloud radars were present. Candidate data sets include, SHEBA, BOREAS, and the continuing ARM project data.

## **Surface Exchange**

Currently, surface fluxes of momentum, sensible heat, and latent heat are parameterized in the models by applying Monin-Obukov similarity theory to the surface layer.<sup>6</sup> These parameterizations use instantaneous values of a prognostic variable from the dynamical portion of the model as input, and various transfer coefficients, roughness lengths, etc. as parameters. Thus, they fail to capture any causal structure that may be present in surface exchange that depends on temporal variability in the data.

ETL has several data sets suited for exploring the possibility of useful temporal patterns existing in surface exchange processes. They include surface flux data from the SCOPE and COPE ocean sensing projects, the SHEBA project in the arctic, and the BOREAS project over forest canopies.

## **Atmospheric Boundary Layer Processes**

Turbulent transport processes in the boundary layer that are most likely to benefit from epsilon machine modeling are the so-called non-local transport processes associated with large-scale eddies and processes that take place at the top of the convective boundary layer. Currently, such processes are still characterized with an eddy diffusivity.<sup>6</sup> Recent analysis of the complexity of vertical wind profiles near the top of the convective boundary layer illustrates that there may be dynamical states of the vertical wind that can be better captured by epsilon machines than by an eddy diffusivity.<sup>7</sup> Suitable ETL data sets for constructing epsilon machine sub-grid models for turbulent transport of momentum and heat in this region of the atmosphere

are profiler and RASS time series data that reach the top of the boundary layer. There have also been occasions when smaller scattering elements are present throughout this region enabling Doppler lidar and radar measurements to reveal the complex behavior of this region.

## 6. Summary

In this memorandum, I first pointed to some glaring deficiencies of present parameterization frameworks for representing physical processes that take place on a scale smaller than the grid size of a climate or weather model. In general these processes are complex, i.e. they involve feedbacks and couplings with degrees of freedom that are governed by both the dynamical and sub-grid portions of the model and cannot be fully captured by parameterizations. Also, the existing parameterization framework makes no attempt to optimize model performance for a fixed computational memory. I referred to how the epsilon machine approach to modeling complex systems can overcome both of these deficiencies by *discovering* rather than imprinting causal pattern and required memory in experimental data. I proposed a specific procedure for constructing epsilon machine sub-grid models beginning with specific partitioning requirements for the raw data and ending with a proposed validation of the method based on single column model performance. Finally, I listed three categories of sub-grid processes that could be modeled with epsilon machines constructed from existing data on hand at this laboratory.

## References

1. Crutchfield, J.P. and Young, K. Inferring statistical complexity, *Phys.Rev.Lett.* **63**, 105 (1989).
2. Crutchfield, J. P. "Knowledge and Meaning: Chaos and Complexity" in Modeling Complex Systems, L. Lam and H.C. Morris, editors, Springer-Verlag, Berlin pp. 66-101 (1992).
3. Kolmogorov, A.N. A new metric invariant of transient dynamical systems and automorphisms in lebesgue spaces" Dokl. Akad. Nauk. SSSR **119**, 861 (1958).
4. Young, Karl (private communication, 1999).
5. [Http://www.cgd.ucar.edu/cms/sccm/userguide.html](http://www.cgd.ucar.edu/cms/sccm/userguide.html)

6. Kiehl, J.T. et.al., Description of the NCAR Community Climate Model (CCM3), NCAR Technical Note (sept. 1996).
7. Palmer, A. J., Fairall, C. F. and Brewer, W. A. Complexity in the atmosphere, J. Geoscience and Remote Sensing. (Submitted, 1999).