GFD 2012 Lecture 5: Applications of coherent structures to the study of weather and climate

Jeffrey B. Weiss; notes by Srikanth Toppaladoddi & Cédric Beaume

June 22, 2012

1 History of computational weather forecasting

Weather forecast plays an important role in preventing disasters. Weather and climate modeling started approximately one century ago, in 1922 with L. F. Richardson. He divided a region into a grid of cells and did 6 weeks of hand calculations to try and model the pressure. The use of computers in weather prediction started only in 1950, when J. G. Charney and his group completed a two-dimensional weather model and ran it on the Electronic Numerical Integrator And Computer (ENIAC). This early work paved way for the founding of the Geophysical Fluid Dynamics Laboratory (GFDL) in the National Oceanic and Atmospheric Administration (NOAA) to study the physical processes that govern the behavior of the atmosphere and the oceans as complex fluid systems. Computers enhanced numerical modeling of the atmosphere and in 1956 N. Phillips developed a mathematical model to depict monthly and seasonal patterns in the troposphere [2]. This model became the first realistic and successful climate model. In 1963, motivated by the study of atmospheric convection, E. N. Lorenz derived simplified equations of convection rolls and implemented them in a simple program. Computations of the resulting equations led to the discovery of chaotic dynamics [3]. Climate modeling has improved a lot since then, and in 1974 S. H. Schneider & R. E. Dickinson reviewed the advances in the field, stating "climate modeling has possibly now reached a threshold where further progress will lead to potential human benefits" [4].

2 Coherent structures in weather and climate

Despite the progress made since the 70's, weather forecast is an extremely difficult task to do accurately. The dynamics of most flows in the atmosphere and oceans is chaotic and even small perturbations can cause large changes. An idea to improve current forecasts is to use coherent structures as the backbone for geophysical turbulent flows. Although being subjective, these structures can be used to reduce variable description of turbulence. Here, we give an example of such coherent structures and their use in numerical weather prediction.

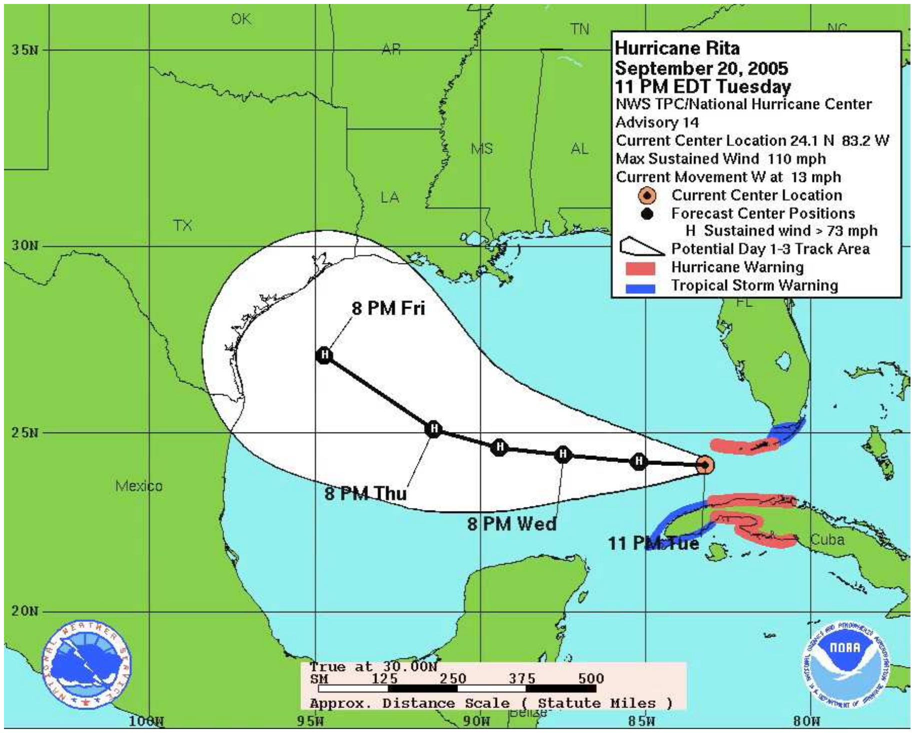

Figure 1: Initial prediction on September 20, 2005 by NWS TPC/National Hurricane Center. Rita is at the tip of Florida and heads west. The black line indicates prediction and the white cone the error. After [1].

2.1 Coherent structures: hurricane Rita

On September 18, 2005 hurricane Rita formed near the Bahamas and became the fourth most intense Atlantic hurricane ever recorded. The potential danger of such an event motivated weather forecasters to predict Rita's trajectory and prompted mass evacuation in coastal Texas. Approximately 3 million people fled prior to Rita's landfall, and the losses were heavy: approximately 100 people died and the damage cost was evaluated at \$ 12 billion. Predictions are reported in figure 1. The predictions contain a large error cone indicating the prediction uncertainty, and are to be compared with Rita's actual trajectory, reported in figure 2. The hurricane finally hit the US at the boundary between Texas and Louisiana 4 days after the initial prediction in figure 1. Rita's trajectory is located within the large cone of errors but very close to its boundary which led to inappropriate decisions in several areas. Although the predictions were rather good, there is room for improvement which translates in more security and less damage.

2.2 Climatic variations: El Niño

Climate and local events are influenced by many variations occurring on different timescales. Among these variations one can cite the North Atlantic Oscillation which consists of atmo-

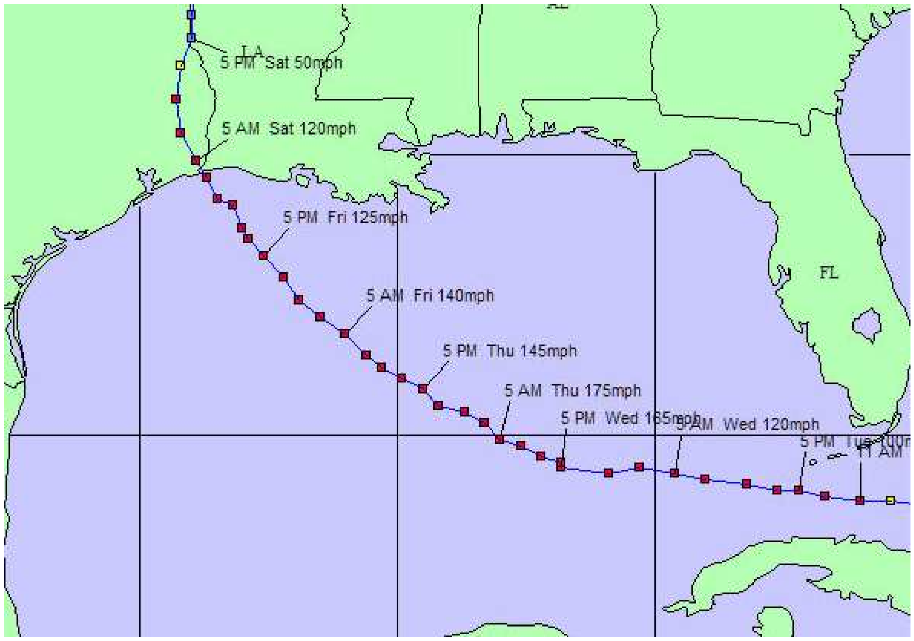

Figure 2: Rita's trajectory. Blue dots indicate low speed (>39mph), yellow medium speed (>58mph) and red fast speed (>74mph). Rita finally hit the boundary between Texas and Louisiana, slightly to the right of the initial predictions 4 days later. After [1].

spheric surface pressure oscillations between the Icelandic low and the Azores high. Another important example of climate variability is given by El-Niño. It is a coupled atmosphere/ocean phenomenon characterized by unusually warm ocean temperatures in the equatorial Pacific that has important consequences on the weather around the globe. This phenomenon can be identified in Pacific Sea Surface Temperature (SST) representations. In figure 3 are reported Pacific SST from 1986 to 2007, time increasing downwards. Indonesia is towards the left of the figure while South America is towards the right. The blue areas on the right of the first plot indicate cool water. The temperature of the water in these areas varies seasonally, being warmest in the northern hemisphere springtime and coolest in the northern hemisphere fall. The red areas on the left indicate hot water, usually seen in the western Pacific. El Niño is an exaggeration of the usual seasonal cycle and can easily be identified in the anomalies on the right figure. Indeed, several El Niños can be seen, for example, in 1986–1987, in 1991–1992 and in 1997–1998. These climate patterns cause extreme weather in many regions of the world such as floods and droughts that can affect many countries.

3 Using coherent structures to improve forecasts

In this part, we discuss traditional methods to predict climate change and weather forecast, and introduce how coherent structures could be used to improve forecasts.

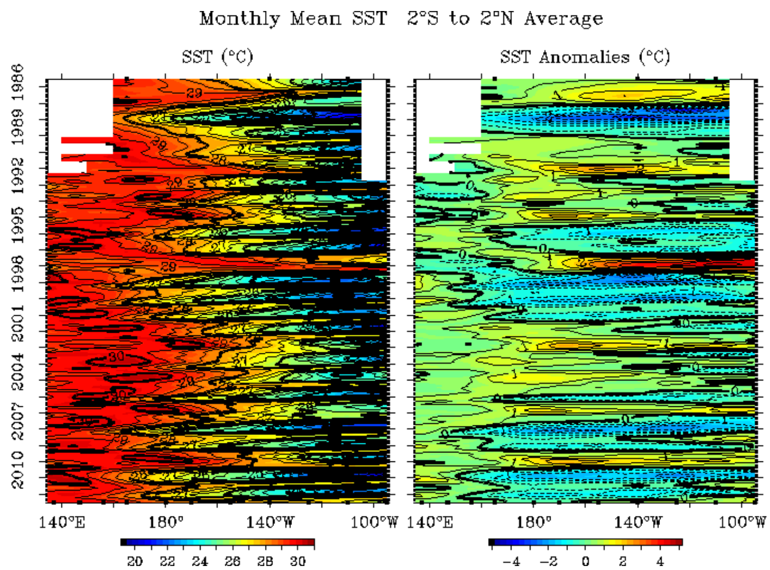

Figure 3: Left: Plot of the tropical Pacific Sea Surface Temperature (SST) in the horizontal, Indonesia is towards the left and South America towards the left. Time is increasing from top to bottom. Red areas indicate hot water while blue ones indicate cool water. Right: Same type of plot but SST anomalies are plotted instead of SST. The anomalies are the difference between the SST and the average for each month.

3.1 Traditional method

The basic method used in numerical weather prediction can be divided into the following steps. A preliminary step consists in covering the area in which the weather is to be predicted with an appropriate model grid. Then, we need to gather information about the current state of the atmosphere (temperature, pressure, wind velocity, humidity, precipitation, etc.). This information is then used to describe the initial condition of the model through a data assimilation scheme which merges the observations with previous model forecasts. As we obtain more observations the initial condition becomes more accurate. Once the initial condition is set, the simulation is run forward to the forecast time. For subsequent forecasts, former predictions are compared to the corresponding observations to verify the model did a good job. Corrections and improvements can then be undertaken. In practice, about half of the cost of numerical weather prediction comes from obtaining observations and the other half comes from running models.

3.1.1 Observation

The first major step in predicting weather is observation. This is done using different kinds of instruments to measure the state of the nature. Ground station are stationary and are used to measure local quantities such as temperature, pressure, wind velocity, precipitation, humidity, etc. Ground stations are often well equipped and very reliable. Observations are also obtain from other instruments such as balloons, satellites, radar, ships or aircraft.

Observations are not typically taken at the same locations as model gridpoints and can be at different times than model timesteps. Data assilimation is used merge these observations onto the model grid.

3.1.2 Data assimilation

Data assimilation is a mature field with much previous work and a sometimes dense nomenclature (see, e.g.,[5, 6]). Different types of data assimilation algorithms are available such as sequential ones (nudging, optimal interpolation, or Kalman filtering), variational ones (minimizing a cost function or 3D variational assimilation) and hybrid ones that combine both methods. Let us introduce a nomenclature of the main state vectors that we will use in the following:

- the true state of the atmosphere $\mathbf{x_t}$. This is a quantity we cannot access.
- the background state $\mathbf{x_b}$ which results from a previous model forecast.
- the observations taken from measurements of the true state $\mathbf{y}$. This quantity is a post-processing of the data collected during observation.
- the analysis $\mathbf{x_a}$. This is the best estimate of the true state obtained by data assimilation.

We give now the example of a 3D variational assimilation (3Dvar). Let us define the observational operator H so that $H(\mathbf{x})$ is the observation obtained from state $\mathbf{x}$. Then we define two error covariance matrices: the error in the background state,

$$B =< (\mathbf{x_b} - \mathbf{x_t})(\mathbf{x_b} - \mathbf{x_t})^T >, \tag{1}$$

and the error in observations,

$$R =< (\mathbf{y} - H(\mathbf{x_t}))(\mathbf{y} - H(\mathbf{x_t}))^T >, \tag{2}$$

where $< \cdot >$ denotes time averaging and T denotes the transpose. Note that these error matrices are difficult to estimate because they involve $\mathbf{x_t}$ which is the unknown true state of the atmosphere. It is commonly assumed that H is a linear operator, $H(\mathbf{x}) = \mathbf{Hx}$ with $\mathbf{H}$ the corresponding matrix. We also make the assumptions that the error is unbiased (the mean is zero) and that the observations and background errors are uncorrelated. We define then a cost function

$$J(\mathbf{x}) = (\mathbf{x_b} - \mathbf{x})^T B^{-1}(\mathbf{x_b} - \mathbf{x}) + (\mathbf{y} - \mathbf{Hx})^T R^{-1}(\mathbf{y} - \mathbf{Hx}). \tag{3}$$

The analysis state $\mathbf{x_a}$ is then by definition the solution that minimizes J, given by

$$\mathbf{x_a} = \mathbf{x_b} + \mathbf{K}(\mathbf{y} - \mathbf{Hx_b}), \qquad \mathbf{K} = \mathbf{BH}^T(\mathbf{HBH}^T + \mathbf{R})^{-1}. \tag{4}$$

Implementing this solution is difficult for two main reasons. First, as mentioned above, obtaining estimates of B and R are difficult. Second, the vectors are very high dimensional (often 10^7-d for $\mathbf{x}$ and 10^5-d for $\mathbf{y}$) making the computations expensive. There are many strategies for dealing with these issues which will not be discussed here.

While numerical weather prediction has seen significant improvements over the decades, there are still failures. Often the failures involve errors in prediction the location of structures such as storms, hurricanes, or the jet stream [7].

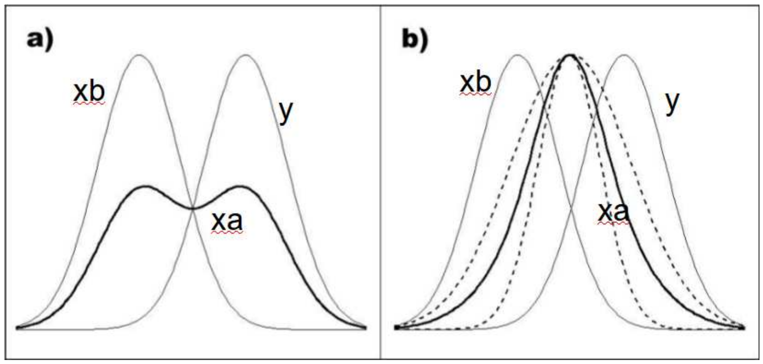

Figure 4: Depiction of a standard assimilation technique applied to a structure with different locations in the observations y and background x_b (a) and a suggested improvement (b). The simplified assimilation technique consists of taking the average between the background and the observation which smears the original structure. The proposed improvement consists of constructing a structure at the mean location of the background structure and the observed structure.

3.2 Improving data assimilation

One drawback of standard assimilation techniques is that they do not preserve the coherence of physically localized structures. As a simple example, imagine an assimilation scheme that takes the average between the observation y and the background x_b. Application of such an algorithm is depicted in figure 4.a. It consists in setting the analysis x_a as the average between the observation y and the background x_b. As the figure shows it, such an algorithm leads to a structure that does not possess the same properties as the observed or background ones: its amplitude is low, its center is no longer a maximum and it can possess two local maxima depending on the position of y and x_b.

To avoid such issues, we desire techniques that can work in many scenarios and applicable to a variety of structures. We refer to methods that explicitly include properties of structures in the data assimilation as "structure assimilation". We build this type of methods in three modular steps.

First, structures need to be identified. This can be done using a variety of existing techniques including those based on wavelets, manifolds, a subjective technique. Once identified, state variables such as the position, size and strength of the structure are defined. The identification is applied to both the background and observations, resulting in background and observed structure variables.

The second step of the method consists in assimilating the structure variables. One can use any of the various data assimilation methods for this step. The result is an analyzed structure that is the best estimate based on the observed and background structure. For example, if the assimilation scheme is a simple average, then the center of the analyzed structure x_a^0 is then given by

$$x_a^0 = \frac{x_b^0 + y^0}{2} \tag{5}$$

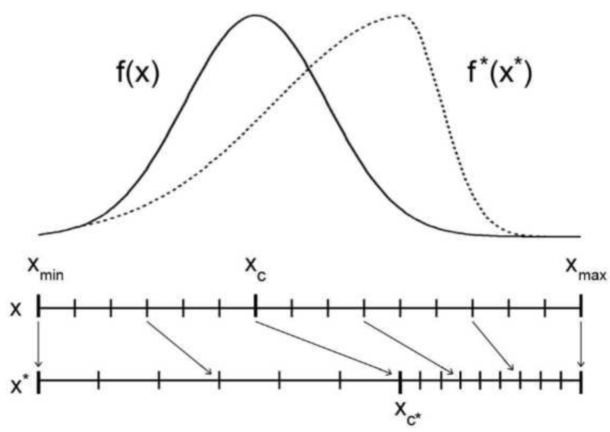

Figure 5: Depiction of a 1D grid morphing technique from a structure f to f^*. The starting structure lies on the grid x. Segments of the grid are stretched or compressed to obtain the grid x^*. The displacement of the meshpoints provides a new structure $f^*(x^*)$.

with x_b^0 being the center of the background structure and y^0 that of the observed one. This type of method is depicted in figure 4.b.

Finally, the last step is the computation of the analysis field on the model grid. While the first two steps implement on well developed techniques, this step has not been previously explored. We propose using the technique of grid morphing to map the structure variables back onto the model grid. Grid morphing is attractive because it has been studied for image processing in the field of computer science. An example of 1D grid morphing is shown in figure 5. The idea of grid morphing is to displace some meshpoints and conserve the value of the field at these points as they are displaced. The resulting shape is then deformed to fit the desired one.

Preliminary numerical simulations using structure assimilation show significant improvements in forecast error [8]. The method was tested using a two-layer QG channel model. Figure 6 compares the errors made by traditional assimilation and structure assimilation. One can see that the error curve for structure assimilation is significantly below that of traditional assimilation. This trend is strengthened by the standard deviation which is smaller for structure assimilation than for traditional assimilation. It follows that structure assimilation can improve forecasts and lead to more reliable results.

4 Climate variability

Climate dynamics is a topic that has assumed importance in recent times due to anthropogenic greenhouse gas emissions leading to climate change. The climate system is a highly nonlinear and a highly coupled system with many feedback mechanisms. Studying the climate system thus requires a hierarchy of models, from highly simplified energy balance models to complex Global Climate Models (GCMs). GCMs include as much of the essential physics as is computationally feasible, and parameterize many processes to reduce their computational cost.

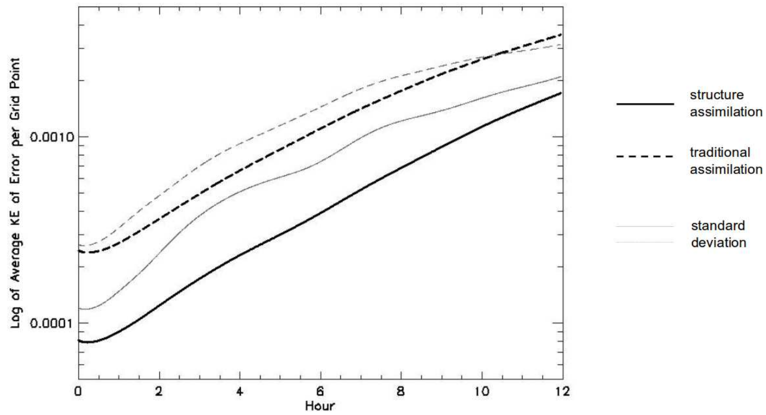

Figure 6: Average kinetic energy of error per grid point as a function of hour after 10,000 forecasts. The straight bold line represents structure assimilation while the dashed bold line represent traditional assimilation. Standard deviations are shown in straight and dashed lines respectively for structure assimilation and traditional assimilation.

In this section, we use stochastic dynamical systems and non-equilibrium statistical mechanics to explore the natural variability of the climate system.

The Earth's climate system has processes whose governing timescales vary from a day to thousands of years, and these nonlinear processes are highly coupled. The natural variability of the climate system takes the form of spatio-temporal patterns which are often difficult to predict. Examples of spatio-temporal patterns include El Niño and the North Atlantic Oscillations.

5 Non-equilibrium Thermodynamics

Non-equilibrium thermodynamics deals with systems which are far from thermodynamic equilibrium. Such systems sustain heat fluxes and produce entropy. A non-equilibrium steady state (NSS) is a statistically steady-state that is kept away from equilibrium by external forcing. One of the simplest examples of this forcing is a system connected to thermal reservoirs at two different temperatures. The NSS then carries a heat from the hot reservoir to the cold reservoir.

A NSS sustains statistically stationary fluctuations about its mean state. Associated with these fluctuations is an entropy production, which can be either positive (entropy producing) or negative (entropy reducing). The Fluctuation Theorem (FT) gives the ratio of the probability of finding fluctuations which increase or decrease entropy. If the probability of a finite time fluctuation changing entropy by S is $P(S)$, and the probability of it changing by the opposite amount is $P(-S)$, then the FT tells us that:

$$\frac{P(S)}{P(-S)} = e^{St}. \tag{6}$$

Equation (6) means that, for a given positive entropy production $|S|$, the probability of a system having an entropy reducing fluctuation is exponentially smaller than the probability of having an entropy producing fluctuation. Because the second law of thermodynamics says that entropy must increase on average, we only expect to see negative entropy fluctuations in thermodynamically small systems on thermodynamically short timescales.

6 Linear Gaussian models

Linear Gaussian stochastic models are some of the simplest models that capture a nonequilibrium steady-state. They also have been effectively used to model a number of phenomena in the climate system. These models take the form

$$\frac{d\vec{X}}{dt} = A\vec{X} + F\vec{\zeta}, \tag{7}$$

where $\vec{X}$ represents the state space of the system. The first term on the right hand side is the linear deterministic dynamics and for the model to remain finite the matrix A must be stable, i.e. the real parts of its eigenvalues are all negative. The second term on the right hand side is additive Gaussian noise where $\vec{\zeta}$ is Gaussian white noise with

$$< \vec{\zeta}(t)\vec{\zeta}^T(t') >= I\delta(t-t') >, \tag{8}$$

where T denotes the matrix transpose and I is the identity matrix. The diffusion matrix which characterizes the noise process is $D = FF^T/2$.

The most common approach for constructing linear Gaussian models for climate phenomena is to build empirical models. In the this approach one fits the matrices A and D to data from either observations or numerical models. The data is typically reduced to $O(10-50)$ degrees of freedom through the use of empirical orthogonal functions (EOF). For example, to study El-Niño one uses observations of sea surface temperature and the state space $\vec{X}$ represents the amplitudes of the EOF patterns.

These simple stochastic models often perform surprisingly well when compared to complex dynamical systems models as can be seen from figure 7. However, it is still unclear why the models perform well for some phenomena and not for others.

Much of the recent work on stochastic models in the climate system has focused on the non-normality of the deterministic operator and the amplification of the noise. The property of non-normality is unsatisfying in that the matrix A can be made normal by a suitable coordinate transformation. However, there is a related coordinate invariant property of the system: the violation of detailed balance. Thermodynamically, systems in thermal equilibrium satisfy detailed balance, while systems in a NSS violate detaled balance. A linear Gaussian model violates detaled balance when $AD - DA^T \neq 0$, and produces noise amplification regardless of the coordinate system.

To analyze the time series data, we make us of the stochastic entropy production which is defined as:

$$S = ln\left(\frac{P(X)}{P(X')}\right), \tag{9}$$

where X is a finite time trajectory segment, X' is the time reversed trajectory segment, and $P(X)$ is the probability of finding the segment in a long time series. The stochastic

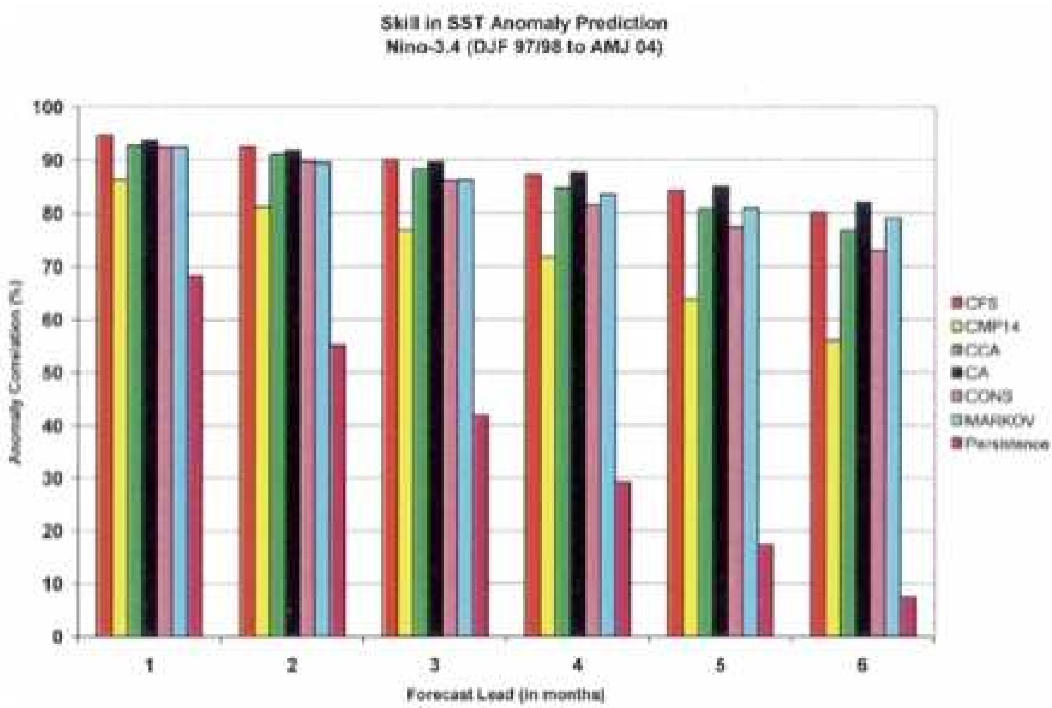

Figure 7: Comparison of different models used to forecast the Nino-3.4 sea surface temperature anomaly. The blue bar is from a stochastic model. From [9]

entropy production requires the trajectory segment over all times between the endpoints of the segment. To apply this to discrete-time climate data, we use a coarse grained entropy, based only on the state vector at the endpoints of the trajectory segment, the endpoint entropy production. The calculations are done with two different methods: the theoretical method which is based on an analysis of the Equation 7, and the direct method, which is based on constructing a pdf of the entropy production of the individual trajectory segments in the data. Agreement between the two methods demonstrates the self-consistency of the linear Gaussian model applied to the data. Computation of the entropy production in a linear Gaussian model of tropical SST shows that we observe negative entropy producing fluctuations on timescales of months (Figure 8). This demonstrates that tropical SST dynamics on monthly timescales is a thermodynamically small system.

In summary, many aspects of natural climate variability takes the form of well-defined patterns. Natural climate variability has a large human impact, and is often poorly captured by GCMs. Climate variability can be modeled as fluctuations about a nonequilibrium steady state in a thermodynamically small system. This suggests that improved understanding of nonequilibrium steady states could have a significant impact on understanding the climate system.

References

[1] http://www.nws.noaa.gov/rita.

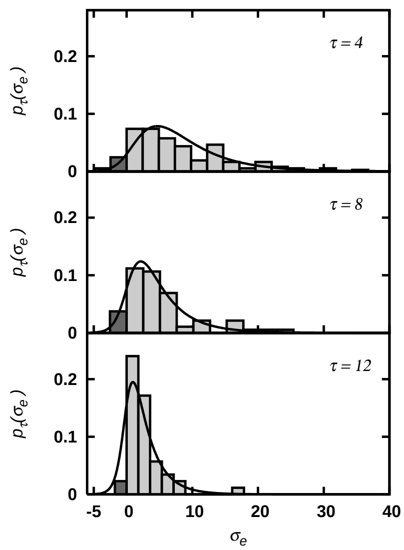

Figure 8: Probability distribution function of the endpoint entropy production. The solid curve is the distribution from the theoretical method and the binned distribution is calculated using the direct method. The darker shaded bins show trajectory segments with negative entropy production. From [10]

[2] N. A. Phillips, *The general circulation of the atmosphere: a numerical experiment*, Royal Meteorological Society **82** (352): 123–154, 1956.

[3] E. N. Lorenz, *Deterministic nonperiodic flow*, Journal of Atmospheric Sciences 20: 130–141, 1963.

[4] S. H. Schneider & R. E. Dickinson, *Climate modeling*, Reviews of Geophysics and Space Physics 12: 447–493, 1974.

[5] F. Bouttier & P. Courtier, *Data assimilation concepts and methods*, Meteorological Training Course Lecture Series, 1999.

[6] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, 2003.

[7] E. Kalnay, S.J. Lord, R.D. McPherson, Maturity of Operational Numerical Weather Prediction: Medium Range, *Bulletin of the American Meteorological Society* 79, 2753-2769, 1998.

[8] Brad E. Beechler, Jeffrey B. Weiss, Gregory S. Duane, and Joseph Tribbia, Jet Alignment in a 2 Layer Quasi-Geostrophic Channel Using One-Dimensional Grid Warping, *The Journal of the Atmospheric Sciences*, doi: 10.1175/2009JAS3263.1., 2010.

[9] Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. Van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Peña, S. Lord, G. White, W. Ebisuzaki,

P. Peng, P. Xie, 2006: The NCEP Climate Forecast System. J. Climate, 19, 3483-3517. doi: http://dx.doi.org/10.1175/JCLI3812.1

[10] J.B. Weiss, 2009: Nonequilibrium Statistical Mechanics of Tropical Sea Surface Temperature Variability, Geophysical Research Letters, 36, L10705, doi:10.1029/2009GL037812.