Prediction of Tropical Rainfall by Local Phase Space Reconstruction*

H. WAELBROECK AND R. LÓPEZ-PEÑA

Instituto de Ciencias Nucleares, UNAM, Mexico

T. MORALES

Centro de Ciencias de la Atmósfera, UNAM, Mexico

F. ZERTUCHE

Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, UNAM, Cuernavaca, Morelos, Mexico

225 October 1993 and 30 March 1994

ABSTRACT

The authors propose a weather prediction model based on a local reconstruction of the dynamics in phase space, using an 11-year dataset from Tlaxcala, Mexico. A vector in phase space corresponds to T consecutive days of data; the best predictions are found for $T = 14$. The prediction for the next day, $\vec{x}_0 \to \mathbf{f}_L(\vec{x}_0)$, is based on a local reconstruction of the dynamical map $\mathbf{f}$ in an η ball centered at $\vec{x}_0$. The high dimensionality of the phase space implies a large optimal value of η, so that the number of points in an η ball is sufficient to reconstruct the local map. The local approximation $\mathbf{f}_L \approx \mathbf{f}$ is therefore not very good and the prediction skill drops off quickly at first, with a timescale of 2 days. On the other hand, the authors find useful skill in the prediction of 10-day rainfall accumulations, which reflects the persistence of weather patterns. The mean-squared error in the prediction of the rainfall anomaly for the year 1992 was 64% of the variance, and the early beginning of the rain season was correctly predicted.

1. Introduction

The possibility of designing nonlinear models for deterministic chaos (Linsay 1991; Abarbanel et al. 1990) has stirred quite a bit of excitement in recent years. In meteorology, it has suggested a new approach to the method of analogues (Van den Dool 1989; Toth 1989), first pioneered by Lorenz (Lorenz 1963, 1969).

The basic idea is to reconstruct the dynamical model from a dataset (Tsonis 1992), which gives the values of one or more variables of the physical system at regular intervals of time. The reconstruction can be done globally, by function fits (Gouesbet 1991; Giona et al. 1991) or neural networks (Principe et al. 1992; Frison 1990), or locally (Farmer and Sidorowich 1987), in which case one reconstructs the dynamics only in the vicinity of the initial data. The local reconstruction promises to be most practical for the relatively high-dimensional chaotic dynamics typically found in weather and climate systems (Ecsedy and Murphy 1992; Galbraith 1992; Alekseev 1992; Tsonis and Elsner 1988, 1989, 1990), with the exception of some low-dimensional climatic attractors, which have been studied successfully using neural networks (Elsner and Tsonis 1992).

Local reconstruction is based on the following procedure: A phase space point is defined from T_e consecutive measurements of the system at regular intervals of time. The dataset, a set of $n \gg T_e$ consecutive measurements, provides $n - T_e - 1$ data points, $\vec{x}_k$, for which the evolution for one time step is known. The data points within a distance η of the initial point $\vec{x}_0$ are used to reconstruct the one-day dynamical map $\vec{x} \to \mathbf{f}(\vec{x})$ in the vicinity of $\vec{x}_0$. The reconstructed map $\mathbf{f}_L(\vec{x}_0)$ can be simply an average of the $\mathbf{f}(\vec{x}_k)$ over the points within a distance η of $\vec{x}_0$ (order zero fit), a linear interpolation, or an order m polynomial fit; the precision increases with the order m, if a sufficient amount of data is available (Farmer and Sidorowich 1988).

In the case of weather forecasting, two problems are encountered. First, the dataset is not always reliable, nor sufficiently long: data gathering projects are usually limited to 10–20 years, or 3650–7300 daily patterns. Second, a local weather system is never strictly isolated: there are external factors that can have a significant influence, such as cold fronts or El Niño.

* This work is supported by CONACyT Grant 400349-5-1714E and by the Association Générale pour la Coopération et le Développement, Belgium.

Corresponding author address: Dr. Henri Waelbroeck, Instituto de Ciencias Nucleares, UNAM, Circuito Exterior, C.U., A. Postal 70-543, 04510 Mexico.

TABLE 1. The first four columns give the prediction accuracy of the optimized model, averaged over the first 10 days of prediction from each of 90 days from June through August 1992, and the comparison with three standards of reference: random pattern selection, persistence, and the average for the current date. The results are given as a percentage of the range of each variable (100% is a perfect prediction). In the last column, the rms error for next-day anomaly predictions of each variable is given, normalized by the rms of the actual anomaly over the 11 years of observations.

(%)	Skill				
	Random selection	Prediction	Persistence	Seasonal average	Rms error next day, season corrected
Vapor tension					
Max	88.8	93.8	93.3	90.0	35.3
Min	88.4	94.4	92.9	94.6	79.1
Dewpoint					
Max	89.9	94.7	93.9	91.7	39.3
Min	83.5	92.9	91.0	93.0	77.7
Relative humidity					
Max	96.9	98.1	97.5	98.0	90.1
Min	88.9	93.2	91.3	93.9	117.9
Atmospheric pressure					
Max	93.2	95.6	95.9	94.6	67.4
Min	92.1	94.9	94.6	93.5	54.7
Cloud cover (oct)	65.1	77.8	73.3	78.5	105.8
Wind					
Dominant direction	76.3	82.2	77.2	82.1	94.4
Dominant speed	92.8	94.7	93.9	94.2	75.4
Max direction	75.3	83.0	76.8	84.1	105.7
Max speed	91.0	93.2	91.3	92.6	79.4
Precipitation	97.3	97.4	96.5	97.5	97.7
Evaporation	84.3	88.1	84.8	87.5	85.6
Temperature					
Max	87.6	93.1	91.8	89.6	40.1
Min	83.9	93.5	91.8	93.7	82.5
Intemperie	82.0	92.5	89.7	91.7	73.8
Insolation (h)	82.4	86.6	84.2	87.5	107.1
Average	86.3	91.6	89.6	91.0	79.4

We are working on a long-term project that aims to develop a meteorological model that is custom designed for tropical "microclimates," enclosed by topographical boundaries. The first phase, which we are reporting in this short note, consists in a straightforward application of a local reconstruction model. Improvements in the use of the available data by applying a higher-order local reconstruction and reducing the noise (Hammel 1990; Sugihara and May 1990; Elsner 1992), and the introduction of external context variables, will not be considered here.

The dataset consists of 19 weather variables measured daily at a single weather station, located near the town of Tlaxcala, Mexico. The measurements were performed without interruption or change of equipment over a span of 4015 days (11 years). Holes and flagrant errors in the dataset amounted to less than 0.5% of the data and were replaced by typical values to minimize their impact. The 19 variables are listed in Table 1, where we report the skill of the optimized model (section 4). The range of each variable was normalized to the unit interval: $x_k^{(i)} \in [0, 1]$ $(i = 1, \cdots, N; k = 1, \cdots, n)$, where $N = 19$ and $n = 4015$. We will use the bold-faced vector notation for a single day of data, $\mathbf{x}_k \in \mathbb{R}^N$, and arrowed vectors for the phase space points $\vec{x}_k = \{\mathbf{x}_k, \mathbf{x}_{k-1}, \cdots, \mathbf{x}_{k-T_e+1}\} \in \mathbb{R}^{N \times T_e}$. A mutual information analysis (Waelbroeck 1994) shows that the mutual dependence of the variables from one day to the next amounts to less than 10% of the information in one day of data. On the other hand, only about half of the 19 variables measured per day are mutually independent.

In section 2, the local reconstruction model is described. The optimization of the parameters of the model is discussed in section 3. Finally, in section 4, we report the results of the application of the optimized model for the year 1992.

2. The prediction model

We will use the normalized Manhattan distance between two data points $\vec{x}_k, \vec{x}_l$ $(k, l = T_e, \cdots, n - 1)$,

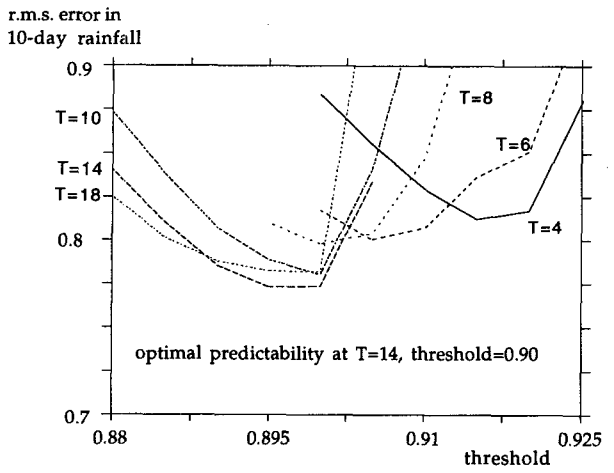

FIG. 1. The optimization of the prediction model using various criteria leads to roughly the same optimum. We give the result of optimizing the mean-squared error on 10-day predictions of rainfall accumulations, a variable which gave the most pronounced optimum. The error is given as a function of the threshold for various values of T. The optimum corresponds to $T = 14$ and a threshold equal to 10% of the range of the variables (which implies that the prediction is a weighted average over roughly 500 points). The large number of points is required for a statistically relevant reconstruction of the local dynamical map in such a large embedding dimension but also implies that the reconstruction is not quite "local," since a significant fraction of all data points lies in the ball of radius 0.1 where the local map is reconstructed.

$$d(\vec{x}_k, \vec{x}_l) = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=0}^{T_e-1}\alpha(t)\,|x_{k-t}^{(i)} - x_{l-t}^{(i)}|, \quad (2.1)$$

where the weights $\alpha(t)$ form a partition of the identity for $t \in \{0, \cdots, T_e - 1\}$, with $\alpha(t)$ decreasing linearly with the delay t so that $\alpha(T_e - 1) = \alpha(0)/2$:

$$\alpha(t) = \frac{4(T_e - 1) - 2t}{3T_e(T_e - 1)}. \quad (2.2)$$

The state of the weather the day after the T_e days in $\vec{x}_k$ is given by $\mathbf{f}(\vec{x}_k) = \mathbf{x}_{k+1}$, where $\mathbf{x}_{k+1} \in \mathbb{R}^N$. We will consider the data points $\vec{x}_k$ such that $d(\vec{x}_k, \vec{x}_0) < \eta$, where η is the radius of the ball around $\vec{x}_0$ where we wish to determine the local map $\mathbf{f}_L$. The reconstructed map is given by

$$\mathbf{f}_L(\vec{x}_0) = \frac{\sum_{k=1}^{n}\lambda(\eta - d(\vec{x}_k, \vec{x}_0))\mathbf{x}_{k+1}}{\sum_{k=1}^{n}\lambda(\eta - d(\vec{x}_k, \vec{x}_0))}, \quad (2.3)$$

where λ is a threshold linear weight function,

$$\lambda(x) = x\theta(x). \quad (2.4)$$

With this choice of weights, one easily shows that the reconstructed map approximates a local linearization if the number of points that pass the threshold is large and their distribution is isotropic (θ is the step function).

An exact linear reconstruction of the dynamics by linear regression was ruled out because of the large embedding dimension: for the optimum value $T_e = 14$ (section 3) one has the very large embedding dimension $d_e = 19 \times 14$, and the linear regression would require inverting a matrix at least this large on each side. For the same reason, the autoregressive models are too slow to be practical. The model presented here can be thought of as a computationally practical way to approximate a local linear regression, and one would expect a similar skill.

The dynamical rule (2.3) takes a weighted average of $\mathbf{x}_{k+1} = \mathbf{f}(\vec{x}_k)$ over the points $\vec{x}_k$ in an η ball centered at $\vec{x}_0$, weighted by $\eta - d(\vec{x}_k, \vec{x}_0)$. For low values of η, this rule gives an associative memory to recall a temporal sequence of patterns, with a storage capacity that can be computed analytically using statistical methods developed for neural networks (Zertuche et al. 1994a,b). Here, we are interested in supercritical values of η, where the learned sequence is unstable and the model acquires positive Liapunov exponents, like the physical system being modeled.

3. Optimization

The prediction model has two parameters, which must be determined by optimization: the scale η at which we reconstruct the dynamics and the number of time steps T_e that define a point in the embedding phase space.

The optimization was performed by removing the last year of data and using it to extract a statistical sampling of initial test points $\vec{x}_0$. Three criteria were used for the optimization. The first two were the overall skill for next-day predictions and the average skill for day-by-day predictions over a span of 10 days. The skill was defined by

$$\text{skill} = 1 - \langle\langle d_1(\mathbf{f}_L(\vec{x}_k), \mathbf{x}_{k+1})\rangle\rangle, \quad (3.1)$$

where

$$d_1(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{N}\sum_{i=1}^{N}|x_1^{(i)} - x_2^{(i)}| \quad (3.2)$$

is the normalized Manhattan distance for a single time step and the average $\langle\langle\ \ \rangle\rangle$ is taken over the test points $\vec{x}_k$ corresponding to the last year of data ($k = n - 365, \cdots, n - 10$). The third criterion was the average mean-squared error of the predicted anomaly of rainfall accumulations over a 10-day period. This last criterion was selected arbitrarily among several variables that have some importance in agricultural planning. The anomaly of a variable at a given date is the difference between the value of the variable and its seasonal average, defined as the average over all available data in a range of plus or minus 10 days around this date, for the 11 years of data (an average over 231 samples). All optimizations gave similar results; we will report

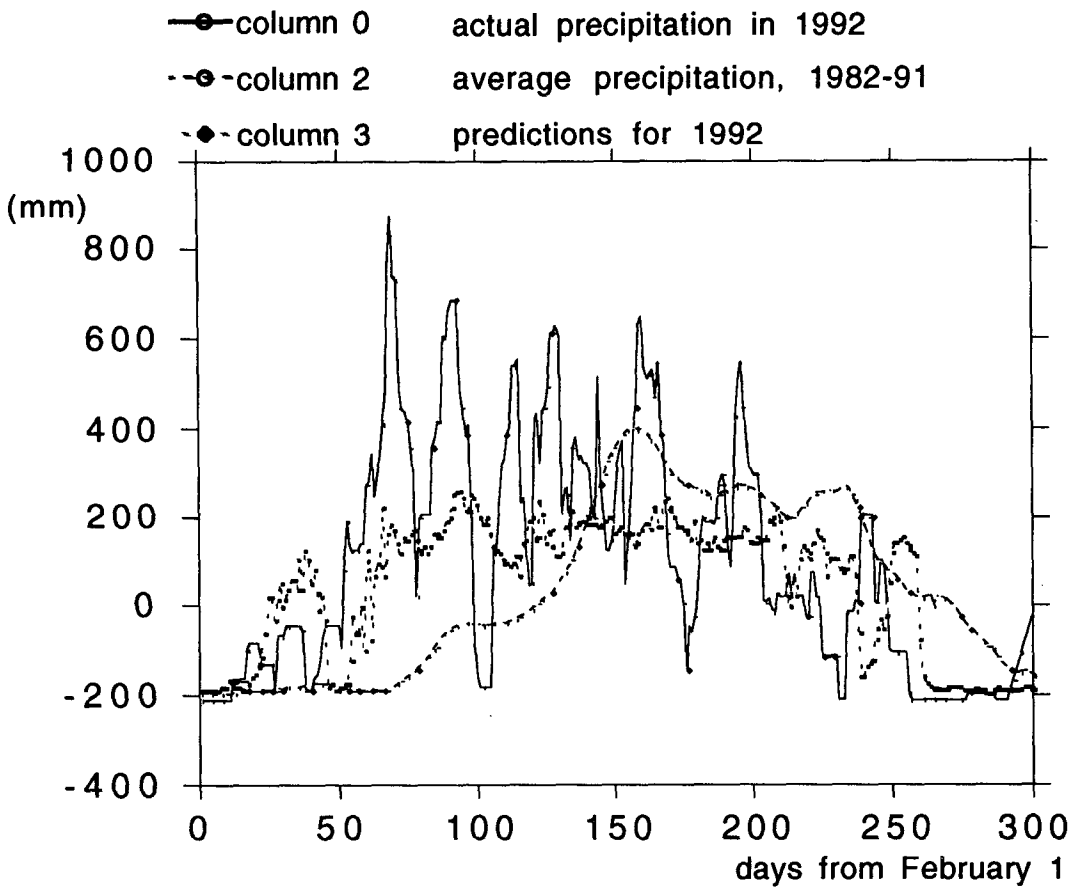

FIG. 2. The actual precipitation curve for 1992, smoothed by 10-day averaging (solid line), is compared to the average for the current date plus or minus 10 days for the span of observations from 1981 to 1992 (thick dotted line) and to the predictions of the model for 10-day rainfall accumulations (fine dotted line). Note that 1992 was an unusual year in that the rain season began almost two months ahead of schedule. This sort of phenomenon is precisely what a model should be able to predict if it is to be useful in practical applications, particularly in agricultural planning. The predictions slightly exaggerate the out-of-season winter rains (first bump at the left) but predict accurately the beginning of the rain season at day 60. The other unusual phenomenon, a dry spell followed by a recovery of rains late in the rain season, is predicted, but with a delay of 10 days.

only the optimization based on the cumulative rainfall predictions.

The results are represented in Fig. 1. The mean-squared error for the predicted anomaly in 10-day cumulative rainfall, relative to the variance of the actual 10-day precipitation anomaly over the 11 years of observation, is given as a function of η for various values of T. The optimum is obtained at $T = 14$, for $\eta = 0.1$. The mean-squared error at the optimum is 0.637, or 63.7%, of the variance. This indicates a mild improvement over the seasonal average, a result which we find encouraging due to the following considerations.

1) Currently available predictions, based largely on general circulation models, are barely competitive with the seasonal average over this timescale (in fact, for that very reason it is the seasonal average most frequently used in tropical regions to recommend planting and harvesting dates).

2) The data can be processed for noise reduction, and one can replace the current "local" interpolation (which is not very good due to the high value of η) by a more sophisticated interpolation scheme that reconstructs the dynamics at a finer scale, down to a limit $\eta = 0.04$, beyond which the external perturbations become relevant (Waelbroeck 1994). Beyond the improvements in the use of the available dataset, the model could also be extended to introduce information on external driving of the local weather system by introducing large-scale predictions from general circulation models.

4. Application of the model

The accuracy of the optimized model is tested along three different criteria: the skill for the prediction of each variable, the mean-squared error of the predicted anomaly of each variable relative to the variance of the

anomaly over the 11 years of observation, and the mean-squared error of the 10-day cumulative precipitation anomaly relative to the corresponding variance. The accuracy according to each of these criteria is compared to three standards of reference: the static model (or "persistence"); the seasonal average, which we defined as the average over all data for the current date plus or minus up to 10 days; and random data selection, where the data for a randomly chosen day in the dataset, $\mathbf{x}_k$, is taken as the prediction.

The average skill of the optimal model ($T = 14$, $\eta = 0.1$) is

$$\max(\text{skill}) = 0.919, \qquad (4.1)$$

which compares to 0.904 for the persistence model, 0.905 for the seasonal average, and 0.865 for random selection. Thus, the model outperforms the seasonal average and persistence for next-day predictions of all weather variables, even though it was optimized for the purpose of predicting 10-day cumulative precipitation. The average skill over 10 consecutive days of prediction is given in Table 1.

The 10-day rainfall predictions for all of 1992 (the last year in the dataset) are compared to the seasonal average and to the observed rainfall accumulations in Fig. 2. Although the predictions do not follow the actual rain profile very closely, the very unusual early beginning of the rain season was predicted by the model.

In summary, we have proposed a weather forecasting model based on a local reconstruction of the dynamics in phase space. The short-term predictability is modest due to the high embedding dimension and the short span of time over which the data were gathered. On the other hand, useful skill was found for longer-term predictions, particularly the 10-day rainfall accumulations, which have some importance in agricultural planning.

A significant improvement of the prediction skill would require improving the use of the available local data to better define the structure of the attractor below the scale $\eta \approx 0.1$ and introducing the external perturbations of the local system from general circulation models. This last task requires a methodology to determine what aspects of large-scale weather patterns are most relevant to the local system—perhaps combining mutual information calculations to select the most important global variables with the application of unsupervised artificial neural networks to categorize the large-scale patterns according to their effect on the local weather system.

For now, we hope that this short note has established that there exists a potential for a new approach to meteorology, based on the reconstruction of the phase space dynamics from data. With this approach, it should be possible to improve on the skill of large-scale models, particularly in the prediction of local, chaotic phenomena.

Acknowledgments. We would like to express our gratitude to the Tlaxcala regional center of the Comisión Nacional del Agua for access to the data, and S. Orozco, J. Jiménez, R. Orea, M. Vázquez, and A. Meneses for the preparation of the dataset. We also gladly express our gratitude to the Comité de Supercómputo of the UNAM for access to the Cray YM-P4. One of us (HW) is also indebted to C. Duqué, director of the Association pour le Développement par la Recherche et l'Action Integrées and the Belgian Ambassador to Mexico for their support of our project.

REFERENCES

Abarbanel, H. D. I., R. Brown, and J. B. Kadtke, 1990: Prediction in chaotic nonlinear systems: Methods for time series with broadband Fourier spectra. *Phys. Rev. A,* **41,** 1782–1807.

Alekseev, V. G., 1992: Analysis of temperature variations from the mid 19th century. *Ian. Fiz. Atm.,* **28,** 926.

Ecsedy, J., and C. G. Murphy, 1992: Global climate warming—Review. *Water Environ. Res.,* **64,** 647.

Elsner, J. B., 1992: Predicting time series using a neural network as a method of distinguishing chaos from noise. *J. Phys. A,* **25,** 843–850.

——, and A. A. Tsonis, 1992: Nonlinear prediction, chaos, and noise. *Bull. Amer. Meteor. Soc.,* **73,** 49–60.

Farmer, J. D., and J. J. Sidorowich, 1987: Predicting chaotic time series. *Phys. Rev. Lett.,* **59,** 845–848.

——, and ——, 1988: Exploiting chaos to predict the future and reduce noise. Los Alamos Preprint LA-UR-88-901.

Frison, T., 1990: Predicting nonlinear and chaotic systems behavior using neural networks. *J. Neural Net. Comp.,* **2,** 45–53.

Galbraith, J. W., 1992: Inference about trends in global temperature data. *Clim. Change,* **22,** 209.

Giona, M., F. Lentini, and V. Cimagalli, 1991: Functional reconstruction and local prediction of chaotic time series. *Phys. Rev. A,* **44,** 3496–3502.

Gouesbet, G., 1991: Reconstruction of the vector fields of continuous dynamical systems from numerical scalar time series. *Phys. Rev. A,* **43,** 5321–5331.

Hammel, S. M., 1990: A noise reduction method for chaotic systems. *Phys. Lett. A,* **148,** 421–428.

Linsay, P. S., 1991: An efficient method of forecasting chaotic time-series using linear interpolation. *Phys. Lett. A,* **153,** 353–356.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.,* **20,** 130–141.

——, 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.,* **26,** 636–646.

Principe, J. C., A. Rathie, and J. M. Kuo, 1992: Prediction of chaotic time series with neural networks and the issue of dynamic modelling. *Int. J. Bifurc. Chaos,* **2,** 989.

Sugihara, G., and R. M. May, 1990: Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature,* **344,** 734–741.

Toth, Z., 1989: Long-range weather forecasting using an analog approach. *J. Climate,* **2,** 594–607.

Tsonis, A. A., 1992: *Chaos. From Theory to Applications.* Plenum Press, 274 pp.

——, and Elsner, J. B., 1988: The weather attractor over very short time scales. *Nature,* **333,** 545–547.

——, and ——, 1989: Chaos, strange attractors, and weather. *Bull. Amer. Meteor. Soc.,* **70,** 14–23.

——, and ——, 1990: Multiple attractors, fractal basins and long-term climate dynamics. *Beitr. Phys. Atmos.,* **63,** 171.

Van den Dool, H. M., 1989: A new look at weather forecasting through analogues. *Mon. Wea. Rev.,* **117,** 2230–2247.

Waelbroeck, H., 1994: Deterministic chaos in a tropical weather system. Mexico Preprint ICN-UNAM-94-03.

Zertuche, F., et al., 1994a: Storage capacity of a neural network with state-dependent synapses. *J. Phys. A,* **27,** 1575–1583.

——, 1994b: Recognition of temporal sequences of patterns using state-dependent synapses. *J. Phys. A,* in press.