MAUSAM, 49, 1 (January 1998), 127-134

551:578.11: 519.246

Dynamic rain model for linear stochastic environments

WALTER RITTER, PEDRO MOSIÑO and ENRIQUE BUENDÍA

Centro de Ciencias de la Atmósfera, Circuito Exterior, CV, 04510, México, DF, México

(Received 25 June 1995, Modified 19 December 1996)

सार — आधुनिक कृषि को विकसित करने के लिए अनिवार्य प्रबंध पर ध्यान देने की आवश्यकता है क्योंकि इससे क्षेत्रीय परिस्थितियों के सन्दर्भ में खाद्यानों के उत्पादन से जुड़े हुए जलवायविक, जैविक, आर्थिक, सामाजिक तथा राजनीतिक कारकों के बीच परस्पर संबंधों की जटिलता का सुव्यवस्थित विश्लेषण किया जा सके।

इसके साथ ही साथ, जलवायविक परिवर्तनों तथा समाज पर पड़ने वाले उनके संभावित प्रभावों के पूर्वानुमान की क्षमताओं को विकसित करने की भी आवश्यकता है। समुचित स्थानीय जलवायु के लिए सुसंगत प्रयोगों द्वारा कृषि पर इसके प्रभाव को कम करना न केवल प्रशंसनीय है बल्कि यह एक आधारभूत आवश्यकता भी है। इसके अतिरिक्त, कृषि तथा मवेशियों के विकास के उत्क्रमण जोखिम को देखते हुए उन क्षेत्रों की और अधिक जानकारी प्राप्त करने में इन अध्ययनों की उपयोगिता की महती आवश्यकता है।

इस शोध पत्र में उपरोक्त लक्ष्यें की प्राप्ति के लिए एक साँख्यिकी मॉडल का उपयोग किया गया है। टलैक्सक्ला राज्य (मैक्सिको) के अनेक क्षेत्रों की वर्षा की विभिन्नता का विश्लेषण अन्तर और आभ्यांतर वार्षिक संबंध दोनों को ध्यान में रखते हुए किया गया है। पहले मामले मेंसंचयी वर्षा लॉजिस्टक कर्व का अनुसरण करती है तथा दूसरे में अरैखिकीय फर्स्ट ऑर्डर प्रसंभाव्य प्रक्रिया का अनुसरण करती है।

ABSTRACT. To develop modern agriculture, a vision of an integral management is required, where the complexity of interactions between climatic, biological, economical, social and political factors involved in the food production must systematically be analyzed in a context of regional conditions.

At the same time, it is necessary to develop the ability to forecast both the climatic variations and their possible impact on society. The minimization of this impact on agriculture through consistent practices adequate to local climates, is not only commendable, but basically necessary, besides, the usefulness of these studies in acquiring a better knowledge of those areas with an inversion risk for agricultural and cattle rising development is high.

In this paper a statistical model is used to accomplish the objectives above mentioned. The rainfall variability in several areas of the Tlaxcala State (Mexico) is analyzed with due regard to both inter- and intra-annual relations, considering that the cumulative rainfall, in the former case, follows a logistic curve and in the latter it follows a linear, first order, stochastic process.

Key words—Forecasting, Stochastic, Logistic, Dry farming, Inversion risk, Mexico.

1. Introduction

The form of a time-series is closely related to its predictability. If in a time-series there is a discernible form, then this series can be used to predict future values. But if there is no consistency in the data form there cannot be predictability, because in an unpredictable sequence there is no cycle or specific period that prevails upon other cycles. Furthermore, it has a uniform distribution of all cyclic components so that there are no specific cyclic forms. It is this non-

TABLE 1

Parameters of the logistic equation, forecasting and probability occurrence for different class intervals in seven localities of Tlaxcala State, Mexico

Place	Class interval (mm)	K (mm)	r (week^{-1})	Probability of occurrence (%)	Annual mean rainfall (mm) K	Initial rainfall (mm) n (to)
Tlaxcala	400-600	537	0.26	6	836	44.5
19°18′ N	601-800	690	0.25	43		31.6
98°14′ W	>801	848	0.26	51		55.7
Tlaxco	400-600	554	0.21	48	683	156.0
19°38′ N	601-800	645	0.22	31		111.0
98°07′ W	>800	769	0.22	19		181.0
Atlangatepec	400-600	562	0.23	34	654	98.3
19°31′ N	601-800	642	0.22	47		95.3
98°12′ W	>801	744	0.23	17		139.6
Cuapiaxtla	401-600	570	0.22	63	609	121.9
19°18′ N	601-800	794	0.21	37		132.0
97°45′ W						
Ixtacuixtla	400-600	553	0.25	38	653	67.3
19°18′ N	601-800	771	0.20	54		112.0
98°15′ W	>801	851	0.23	8		129.0
Apizaco	400-600	532	0.20	12	786	95.4
19°26′ N	601-800	756	0.22	52		169.0
98°09′ W	>801	1078	0.22	36		120.5
Españita	400-600	766	0.23	36	1018	107.9
19°28′ N	601-800	1470	0.25	64		247.0
98°24′ W						

uniform distribution of cyclical components that makes prediction possible.

The predictability of a time-series is described in terms of the autocorrelation function. Thus, to determine whether there is an inherent predictability in such a series or not, it must be found if there is correlation among consecutive data. If there is no autocorrelation the sequence is completely unpredictable. For if there is an autocorrelation, this can be used to make a prediction of the future trend based on past observations. The autocorrelation function ρ (h) is the correlation coefficient between observations separated by h time intervals.

In this paper we develop a statistical model based on both the logistic equations and time series properties, in order to obtain an index value for agriculture and cattle raising development risk.

2. The model

For the simulation and forecast of accumulative weekly mean values of rain in the Tlaxcala state, Mexico, Ritter and Guzman (1991) used a logistic model of the form $\frac{dN}{dt} = rN\left(\frac{N-K}{K}\right)$ with very good results. The equation in its predictive discrete form is:

$$N_{t+1} = N_t\left[1 + r\left(1 - \frac{N_t}{K}\right)\right] \quad (1)$$

where N_t is the observed precipitation in the previous period, r is the rate at which the maximum rain cumulative yearly K value is reached, during the normal rainfall period (Table 1).

The seasonal application of a linear model to a sequence K_t of rainfall anomalies, in which there is an autocorrelation between its consecutive values and considering a stochastic variation in the K_t amounts that could be represented as a first-order autoregressive process is given by:

$$K_t = \lambda K_{t-1} + Z_t (-1 < \lambda < 1) \quad (2)$$

where each Z_t is an independent random variable with zero mean and variance σ_z^2 whose distribution is arbitrary and not necessarily Gaussian. Eqn. (2) tells us that the rainfall event K_t at time t is λ times its value at $t-1$ plus a random component Z_t and where λ controls the predictability of the event. If $\lambda=0$, K_t is simply an independent event equal to the random variable Z_t and in this case K_t is completely unpredictable. But if $\lambda>0$ then some of the K_{t-1} values persist in K_t, *i.e.*, the system keeps memory and then the consecutive correlation occurs.

If $\lambda<0$ there is also predictability, but with a negative consecutive correlation. A negative value of λ indicates an oscillatory environment (Feller 1957).

In general, λ could be thought of as a measure of the speed of recuperation of a disturbance. If the amount K_t takes values less than $\overline{K}$, the mean amount, then a high value of λ indicates a slow recovery of the system. While if K_t takes values greater than $\overline{K}$, then a high value of λ indicates a continued persistence.

Thus, besides controlling the predictability of the K_t event, λ also determines the rate of recovery to normal conditions.

The autocorrelation function for an autoregressive process of first order of the logistic model is given as:

$$\rho_K(h) = \lambda^{|h|} -1 < \lambda < 1; h = -1,0,1,... \quad (3)$$

which is a geometric sequence in λ, Feller (1957). The variance of the event, σ_K^2, is both related to λ and to the random component variance σ_z^2 through:

$$\sigma_K^2 = \frac{\sigma_z^2}{1-\lambda^2} \quad (4)$$

Thus, considering a first-order autoregressive process for K_t a particularly simple description of the environment is obtained, where σ_K^2 is a measure of the annual variability of rainfall and λ a measure of its predictability.

If the K_t process is autoregressive and of first order, the variance σ_n^2 of the weekly rain anomalies (n_t), directly related to the annual rainfall anomalies and variance (σ_k^2) is given by Cox and Miller (1968) and Roughgarden (1979) as:

$$\frac{\sigma_n^2}{\sigma_k^2} = \left(\frac{r}{2-r}\right) \Bigg/ \left[\frac{1+(1-r)\lambda}{1-(1-r)}\right] \quad (5)$$

The autocorrelation function of the intensity of the weekly rain anomalies n_t is given by the following formulae:

$$\rho_n(0) = 1$$

$$\rho_n(h) = \frac{(1-r)\rho_n(h-1)+r \qquad (2-r)\lambda}{1+(1-r)\lambda} \quad (6)$$

Hence:

If $r=1$, then from Eqns. (5) and (6) weekly changes in h at $t+1$ follow exactly, or are a result of annual changes in k at t.

As a consequence, the variability and predictability of n_t is identical to that of k_t and then

$$\sigma_n^2 = \sigma_k^2; \; \rho_n(h) = \rho_k(h) \quad (7)$$

For an unpredictable environment ($\lambda=0$), Eqns. (5) and (6) become

$$\frac{\sigma_n^2}{\sigma_k^2} = \frac{r}{2-r}; \; \rho_n(h) = (1-r)^{|r|} \quad (8)$$

The important thing to note from Eqn. (8) is that, besides the constant σ_k^2, both the variability (σ_n^2) and the predictability of n_t are determined by r and as it tends to zero, n_t becomes less sensitive to the h values. This decreases the variance, σ_n^2 and increases the serial correlation $\rho(h)$ for all h. Thus, in unpredictable environments whatever be the predictability

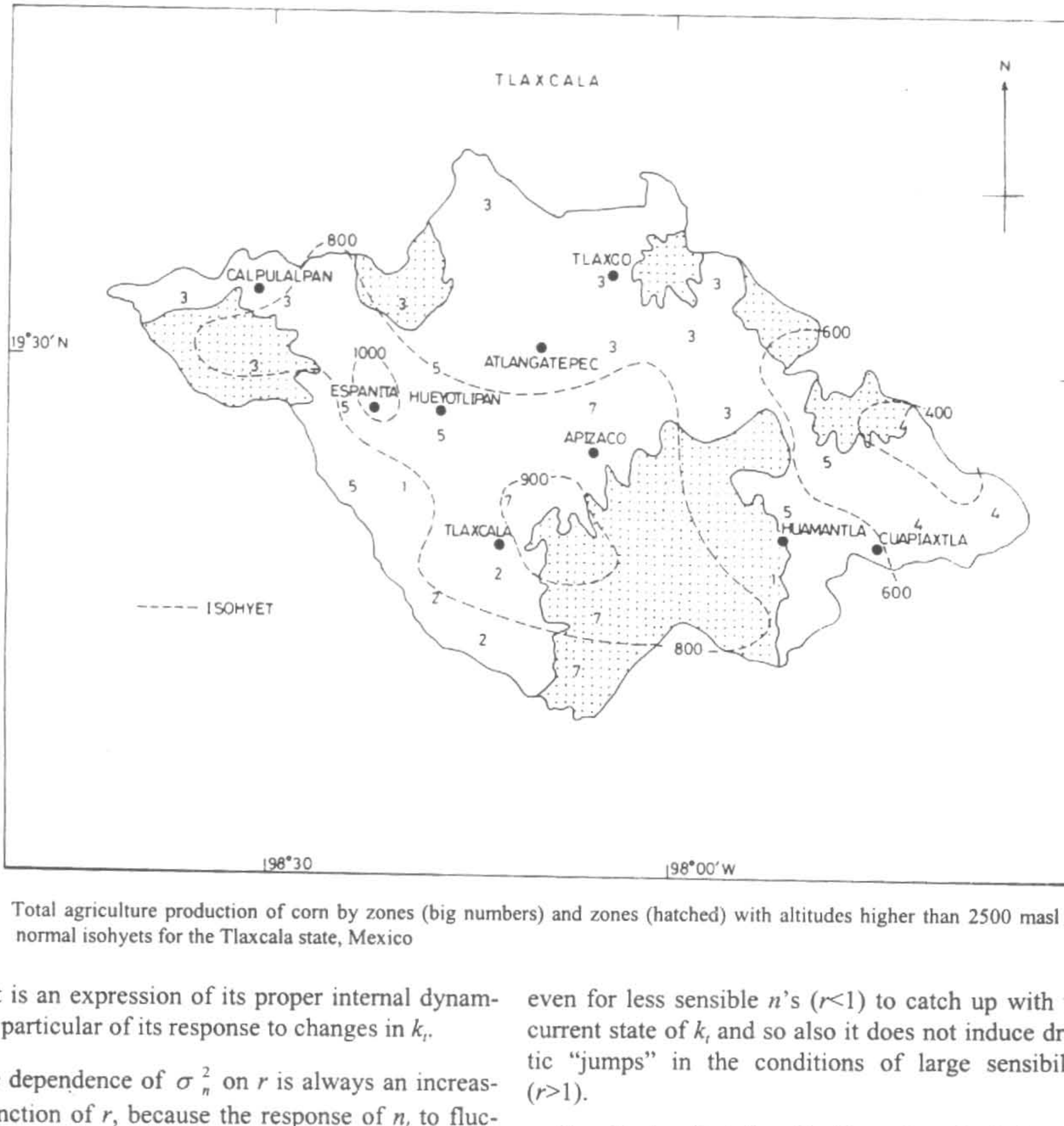

Fig. 1. Total agriculture production of corn by zones (big numbers) and zones (hatched) with altitudes higher than 2500 masl and normal isohyets for the Tlaxcala state, Mexico

in n_t it is an expression of its proper internal dynamics, in particular of its response to changes in k_t.

The dependence of σ_n^2 on r is always an increasing function of r, because the response of n_t to fluctuations in k_t is controlled by r. Augmenting r the response of n_t increases and also σ_n^2. The dependence on λ is because, for large values of λ, σ_n^2 tends to and becomes equal to σ_k^2 over wide ranges of r. High values of λ indicate a strong, positive correlation between consecutive values of the k's. Thus with large values of λ, the changes in k_t between consecutive times are rarely large; instead, the total variations in k_t as a measure of σ_k^2 are realized by a rather slow oscillation in the k_t values.

This fact keeps σ_n^2 close to σ_k^2. The slow course of the variance quotient with a large λ gives time even for less sensible n's ($r<1$) to catch up with the current state of k_t and so also it does not induce drastic "jumps" in the conditions of large sensibility ($r>1$).

Considering first the situation where $\lambda=0$, (unpredictable environment) the autocorrelation is controlled by the degree of response of r only.

If $r=1$, $\rho_n(h)=1$ for $h=0$, and $\rho_n(h)=0$ for $h>0$; $\rho_n(h)$ is identical to $\rho_k(h)$. As r exceeds one, the autocorrelation function oscillates reflecting the peaks and valleys of n.

The situation where $\lambda \neq 0$ are natural extensions of those for which $\lambda=0$.

When $\lambda<0$, the autocorrelation function of n_t acquires an oscillatory character due to the explosive nature of the environment.

Any degree of both variability and predictability of n_t could be produced by many combinations of r and λ. To separate the causes requires continued observation in time.

A measure of the tracking error is the average discrepancy between n_t and k_t defining the relative follow up error as:

$$e = \frac{(k_t - n_t)^2}{\sigma_K^2} \tag{9}$$

This quantity is the tracking mean with respect to the environmental variability of k_t. If it is considered that k_t is given by a first-order autoregressive process it can be found how e depends on both the response sensibility of r and on the environmental predictability λ. Carrying $n_t = \Sigma_{i=1}\, r(1-r)^{i-1}\, k_{t-1}$ into the above mentioned equation it is obtained.

$$e = \frac{\sigma_K^2 - 1\Sigma r(1-r)^{i-1}\gamma_K(i) + \Sigma_{i=1}^{\infty}\Sigma_{j=1}^{\infty} r^2(1-r)^{i-1}\gamma_K(i-j)}{\sigma_K^2} \tag{10}$$

where $\gamma_K(h)$ is the autovariance function of K_t, the equation gives the follow up relative value for any K_t process. In the special case of a first-order autoregressive process for K_t, it results in:

$$\gamma_K(h) = \sigma_K^2\, \lambda^{|h|} \tag{11}$$

which substituted into the previous equation it is finally obtained

$$e = 2(1-\lambda)/[2(1-\lambda)] + r[(3\lambda-1) - r\lambda] \tag{12}$$

which is the tracking relative error for the case of a first-order autoregressive process k_t. Thus, the mean discrepancy between n_t and k_t is a function of the response of r and the environmental predictability λ.

The r that makes the tracking error a minimum, in a series of simulations, is defined as the best r, and this last is found by differenciating Eqn. (12) with respect to r, equating to zero and solving.

The intuitive meaning of these equations is as follows: in an unpredictable environment the weekly rain responding to changes in k_t, it is not likely to better the tracking, because in the next interval it is possible for k_t to have some value very different to that needed to stay near its previous value.

TABLE 2
Range values of corn production (kg/ha)

Number	Equivalent Production
1	> 1200
2	1000-1200
3	800-1000
4	600-800
5	400-600
6	200-400
7	< 200

In this situation, the minimal tracking error arises from a little sensible n_t to k_t fluctuations whose intensity remains fixed in the mean value of K_t. However, if the environment is predictable, in the sense that when changes occur it is likely that they persist, then responding to the changes we will obtain a reduction in the tracking error.

3. Results

Both the 30-year (1960-90) weekly and annual rainfall data used in this paper were obtained from the National Commission for Water, in the Tlaxcala State.

The presence of rain in the Tlaxcala State is strongly influenced by its orography, (Ritter and Guzman 1991), as can be seen in Fig. 1 for isolines of normal precipitation where mean annual maxima are distributed along a curve which goes from the west (Calpulalpan) towards the southern part of the state (Tlaxcala city) through the larger values observed at Españita.

This axis of major annual rainfall amount, coincides with the relatively low terrain of the Tlaxcala State (and higher total production of corn, Table 2) which is visited by either southwesterly storms that may bring heavy rain or northeasterly ones which produce very little rainfall over the area; notwithstanding that most of the air moisture comes from the Gulf of Mexico, on account of the rain-shade effect caused by the high north south, mountain ridges between the lower terrain of the Tlaxcala State and the very lower ground of the Gulf coast.

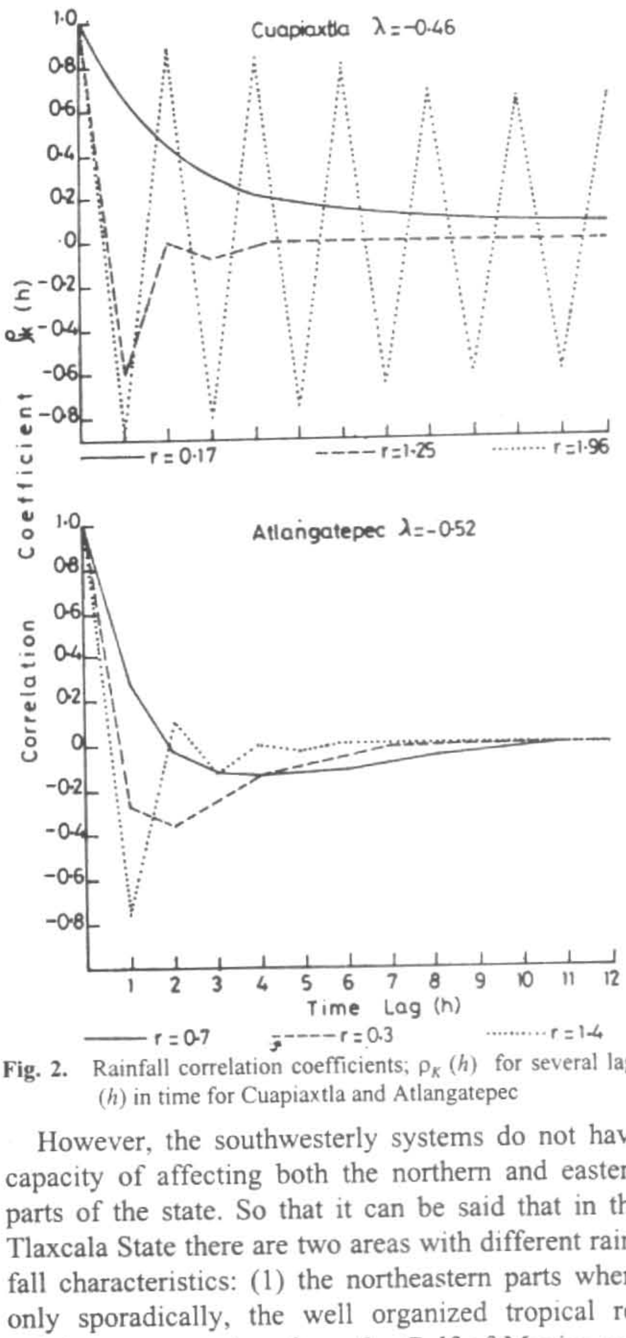

Fig. 2. Rainfall correlation coefficients; ρ_K (h) for several lags (h) in time for Cuapiaxtla and Atlangatepec

However, the southwesterly systems do not have capacity of affecting both the northern and eastern parts of the state. So that it can be said that in the Tlaxcala State there are two areas with different rainfall characteristics: (1) the northeastern parts where only sporadically, the well organized tropical revolving storms arrive, from the Gulf of Mexico and (2) the southwestern part where there arise both the strong weather systems and a lot of small, but more frequent ones, that lead to large amounts of the total annual rainfall.

Under these conditions the rainfall predictability in both areas will be different and will give diverse situations, more random where only the strong systems arrive and more uniform in the areas with the presence of both systems. This situation is shown in applying our stochastic, linear model to Cuapiaxtla and Atlangatepec stations, Fig. 2, which have a λ,

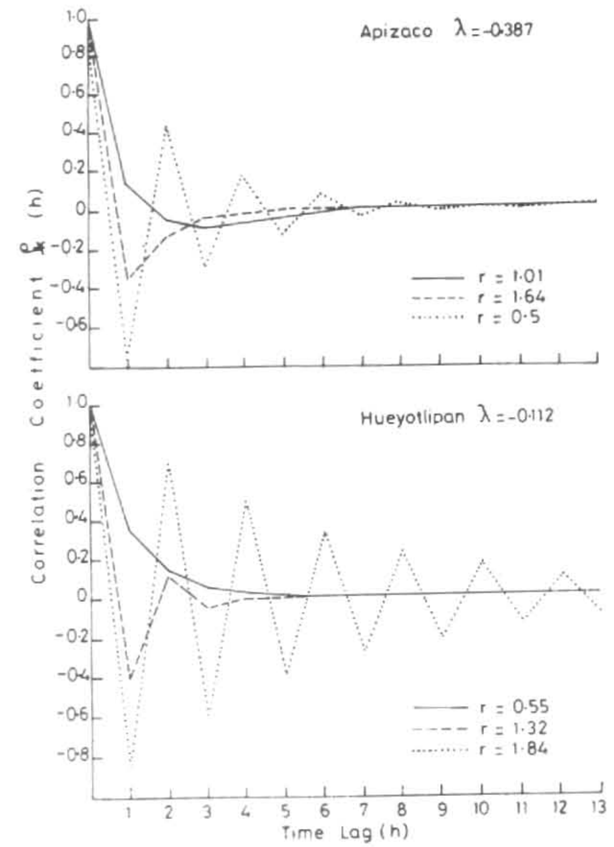

Fig. 3. Rainfall correlation coefficients: ρ_K (h) for different lags (h) in time for Apizaco and Hueyotlipan

with nearly equal and negative values of 0.46 and 0.52 respectively. The fact that Cuapiaxtla is located in the eastern area makes it susceptible to an unstable behaviour, the same as Atlangatepec's in the north. It is worth pointing out that Cuapiaxtla with r values of 1.25 begins to show some instability and that for Atlangatepec r values of 1.40 present even larger manifestations of instability which seems to indicate the continuity of behaviour of one and the same system. Another similar situation is that shown both by Apizaco and Hueyotlipan Fig. 3, located within the 800 mm isohyet where the r-values seem to present situations of the same system. r values of 1.01 begin to show instability in the former, incremented to r=1.32 in the latter. There, r values increase in a proportional way with r=1.64 in Apizaco and finally in Hueyotlipa for r=1.84, the instability becomes manifest for the 13 lags considered.

The average tracking error of n_t and k_t, relative to the environmental variability of K_t, Fig. 4 gives a similar behaviour to that already observed in the previous ratio values of σ_n^2/σ_k^2, being both Tlaxcala

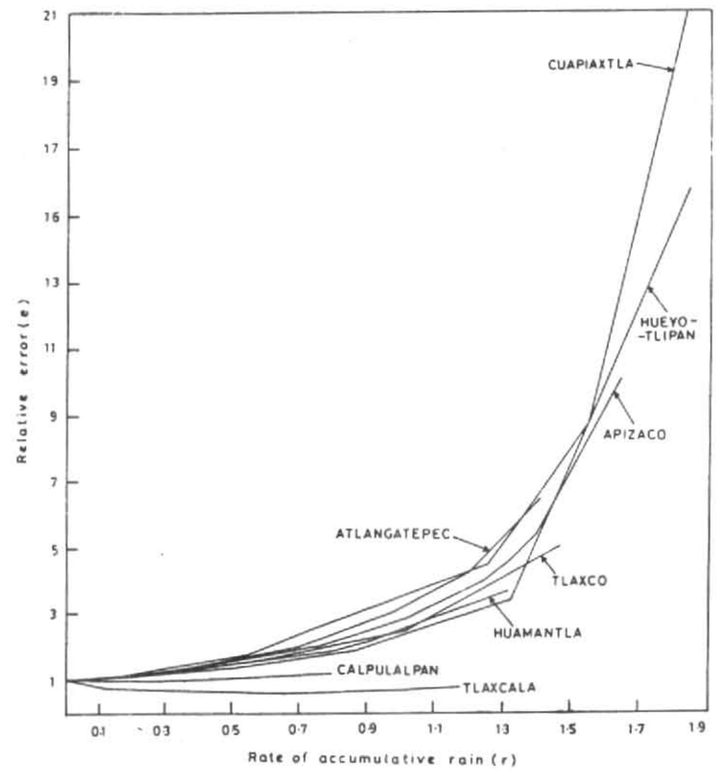

Fig. 4. Tracking relative error (e) of the anomalies of the annual precipitation change with respect to its distribution throughout the year, for different r values, in eight analyzed stations of the Tlaxcala State

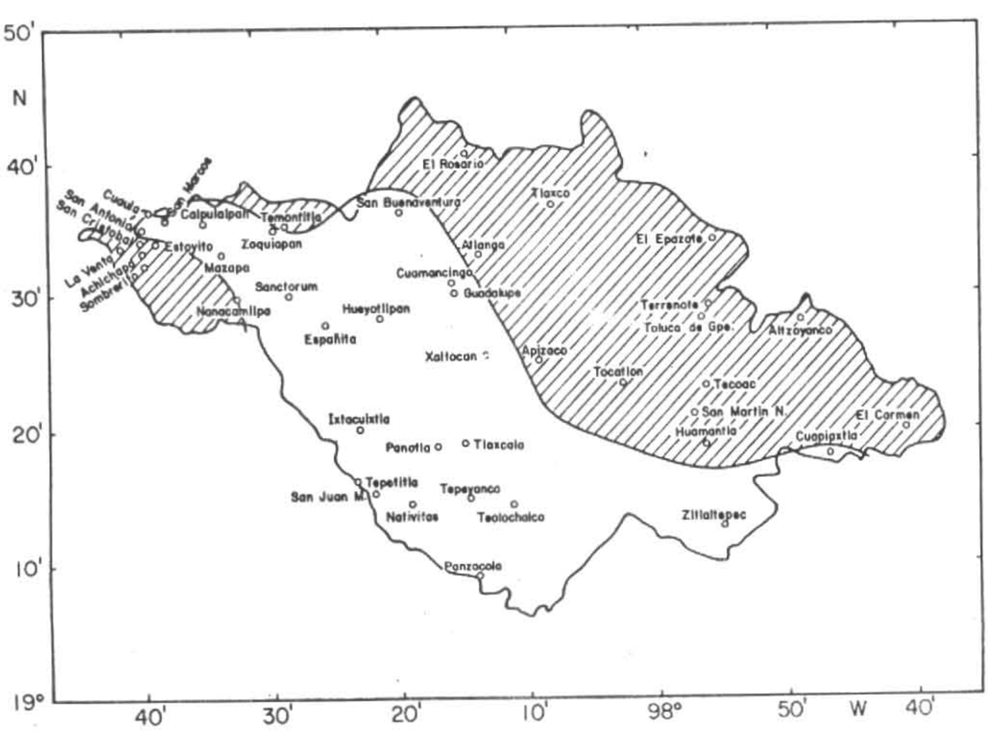

Fig. 5. Zones with higher (hatched) and lower (white) risk for economical inversion in agriculture and cattle rising development

and Calpulalpan those with highest stability and following capacity in its behaviour, followed by Hueyotlipan which presents a great instability for r-values larger than 1.3; followed by Huamantla, Tlaxco and Atlangatepec which are in this respective order more unstable than Hueyotlipan; but for $r<1.3$ they show a great stability.

Cuapiaxtla, the more unstable area, presents large changes when r is higher than 1.2 and 1.5. Apizaco is the area which shows more changes in its tracking trajectory, presenting unstability when r takes values higher than 0.7, 1.0, 1.2 and 1.4. From what has been stated above it can be concluded that to know the rainfall predictability in the Tlaxcala State the stability is analyzed with a linear stochastic model indicating an orographic influence in the area of study.

Thus, there are two distinct areas: an area of low terrain within reach of both strong as well as weak rainfall systems (western and southern parts of the state) and another more unstable area with approach for only strong rainfall systems (northern and eastern parts of the Tlaxcala State). The former being more stable in its behaviour is less risky for inversion and therefore more predictable, than the latter. Within the first area Tlaxcala city and Calpulalpan are good examples of a stable and predictable area, while Cuapiaxtla would belong to the second area, showing the largest instability values, (Fig. 5).

The various analyzed areas present distinct instability degrees with respect to r and Apizaco is the one that presents a larger number of manifestations of changes with respect to this variable.

The stable areas are located in the zones with higher mean annual rainfall. Thus the former shows a shorter and more accentuated rainy season more easy to forecast; because in a larger number of cases, they tend to reproduce the 30-year normal values.

When the rainy season extends itself beyond the normal fading date, the weekly variability of the rain increases and it makes a great difference with respect to the variance of the mean annual precipitation, *i.e.*, if the rainy season prolongs beyond the expected normal terminal time, the weekly observed rains are likely to show great discrepancies with the average annual values.

Acknowledgements

The authors wish to thank the help received from Engineer Alfonso Estrada Betancourt, Rafael Patiño and Ma. Esther Grijalva during the elaboration of this paper.

References

Cox, D.F. and Miller, H.D., 1968, "The Theory of Stochastic Processes", Wiley, New York, 389p.

Feller, W., 1957, "An introduction to probability theory and its applications", 2nd. Ed. Vol. 1, John Wiley & Sons, New York, 346p.

Ritter, O.W. and Guzmán, S.R., 1991, "Simulación y pronóstico de lluvias, una aplicación al estado de Tlaxcala, (México), *Geofisica Internacional*, **Vol. 30**, 183-192.

Roughgarden, J., 1979, "A simple model for population dynamics in stochastic environments", *Am. Nat.*, 713-736.