Conditional Distributions of Neyman-Scott Models for Storm Arrivals and Their Use in Irrigation Scheduling

JORGE A. RAMIREZ AND RAFAEL L. BRAS

Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge

This paper solves the deficit irrigation scheduling problem assuming that rainfall arrivals obey a Neyman-Scott cluster model. The implied dependence between storms is represented by derived conditional distributions of the occurrence of rainfall based on the history of past storm arrivals. This new result is used on a physicostochastic model of the soil moisture history which in turn leads to an optimal control algorithm for making irrigation decisions. The decision of when and how much to irrigate now depends on measured soil moisture, available irrigation water, and time since the last rainfall occurrence.

INTRODUCTION

Recently, *Córdova and Bras* [1981] and *Bras and Córdova* [1981] proposed a new solution to the irrigation scheduling problem in regions of deficit irrigation. The novel ideas of the formulation were the analytical combination of stochastic models of rainfall arrivals with physically based models of the soil-plant system. A stochastic model of soil moisture which depended on climatic probabilistic parameters and on physical description of the soil-plant system was thus derived.

In their work, *Córdova and Bras* assumed that storm arrivals were independent Poisson events. An alternative is to account for the widely recognized, but often neglected, statistical dependence in storm arrivals [*Gabriel and Newmann*, 1957, 1962; *Smith and Schreiber*, 1973; *Kavvas and Delleur*, 1975]. Recently, authors like *Kavvas and Delleur* [1975] and *Gupta and Waymire* [1979] have argued that this dependence is due to clustering of rainfall in time and space. They suggest modeling storm arrivals with cluster point processes like the Neyman-Scott model.

This work uses the concept of cluster models for storm arrivals. The conditional distribution of the occurrence of future events given a history of the timing of past events is derived for the Neyman-Scott model. This new result is then used with a reformulated model of the irrigation scheduling problem. Now the decision of when and how much to irrigate is based on the time since the last observed storm, the effective soil moisture content, and the amount of available irrigation water.

A case study is presented to illustrate the solution. Significantly, there is no important difference between the theoretical net benefits derived from assuming a homogeneous independent Poisson model or a dependent Neyman-Scott (N-S) model during the growing season. A nonhomogeneous Poisson model yields higher theoretical benefits. As expected, the conditional N-S model leads to improved irrigation control as compared to the unconditional N-S model. From this case study, however, no definite conclusion can be drawn regarding whether the Neyman-Scott model better represents the process of storm arrivals as opposed to the Poisson model. However, the expected opportunity loss of choosing the wrong model can be significant.

Interesting implications of the results are that data manipulation can have a significant effect on the ultimate model

chosen; that including the often neglected precipitation dependence structure leads to better irrigation control; that the chosen precipitation model has a profound effect on expected net benefits, thereby affecting resource allocation and managerial decision making; and that dependence commonly attributed to clustering may be due to the low-frequency effects of hard to detect nonhomogeneity in rate of storm arrivals.

THE NEYMAN-SCOTT CLUSTER MODEL: REVIEW

Clustering is one of the most common causes of dependence in natural phenomena. The spatial distribution of galaxies has been hypothesized to occur in clusters and modeled as such by *Neyman and Scott* [1952]. Cluster dependence structure has also been observed in the occurrence of earthquakes and their aftershocks. *Vere-Jones* [1970] used cluster processes to model earthquake occurrence. In fact, the literature is full of examples of various natural phenomena exhibiting clustering behavior. In the hydrologic literature, works by *Le Cam* [1961], *Kavvas and Delleur* [1975, 1981], *Kavvas* [1982a, b], *Gupta and Waymire* [1979], and *Waymire and Gupta* [1981a, b, c] have popularized the use of cluster models. The uninitiated reader is invited to read the above references for complete details of the results given in this section.

One important class of cluster processes is the so-called Moyal cluster processes [*Moyal*, 1962]. They are the superposition of finite point processes and in the real line can be constructed as follows. Define on the real line, R , a point process $\tilde{N}(\cdot)$ ($N(\cdot)$ is the counting measure of occurrences in subsets of R , i.e., for any subset $A \subset R$, $N(A)$ is the number of events in A), as the process of cluster centers. Associate with each arrival time t_i of $\tilde{N}(\cdot)$ a finite point process $N_i(\cdot | t_i)$. Heuristically, $N_i(\cdot | t_i)$ represents the cluster members associated with the i th cluster. Then, the cluster process is given by

$$N(\cdot) = \sum_{i=1}^{\infty} N_i(\cdot | t_i) \quad (1)$$

Among the members of the class of cluster processes defined by (1) is the Neyman-Scott (N-S) cluster process. The cluster process of (1) becomes a Neyman-Scott cluster process whenever the process of cluster centers $\tilde{N}(\cdot)$ is Poisson and whenever a random number of points $N_i(\cdot | t_i)$ are independently distributed about the cluster centers according to a common distribution. When cluster models of the kind presented above are used for modeling precipitation occurrence in time one effectively assumes that rainfall occurrence is a two-level process. In the primary level, or parent process, there is the occurrence of rainfall generating mechanisms (RGM's), $\tilde{N}(\cdot)$, or

cluster centers. In the secondary level, for each cluster center at t_i , there is the actual associated number of rainfall events, $N_i(t_i)$, or cluster members. The RGM's can be, for example, cyclonic belts persisting over a region, or fronts sweeping over a given area [Kavvas and Delleur, 1975].

The probability generating functional (pgfl) of the N-S cluster process is

$$G[\xi] = \exp \left\{ \mu \int_{-\infty}^{\infty} \left[g_{N_i} \left(\int_{-\infty}^{\infty} \xi(t + \zeta) f_T(\zeta) d\zeta \right) - 1 \right] dt \right\} \quad (2)$$

where g_{N_i} is the probability generating function (PGF) of the random variable representing the number of cluster members (rainfall events) in a cluster centered at time t_i , $f_T(\zeta)$ is the probability density function (pdf) of time positions of the cluster members within each cluster, and μ is the parameter of the Poisson distribution for cluster centers.

Under suitable choices for the function $\xi(\cdot)$, (2) yields the complete finite dimensional structure of N-S cluster process. For example, to obtain the univariate representation of the N-S cluster process and assuming that the PGF of the number of rainfall occurrences $N(0, t_1]$ is desired, define

$$\xi(x) = 1 - (1 - z)I_{(0, t_1]}(x) \quad |z| < 1 \quad (3)$$

where $I(\cdot)$ is the indicator function such that

$$\begin{aligned} I_{(0, t_1]}(s) &= 1 & s \in (0, t_1] \\ I_{(0, t_1]} &= 0 & \text{otherwise} \end{aligned} \quad (4)$$

Substituting (3) into (2) yields

$$g(z) = \exp \left\{ \mu \int_{-\infty}^{t_1} [g_{N_i}(1 - (1 - z)p(t)) - 1] dt \right\} \quad (5)$$

where

$$p(t) = \int_0^{t_1} f_T(\zeta - t) d\zeta \quad (6)$$

Equation (5) is the univariate PGF of the N-S cluster process.

To obtain the complete finite dimensional stochastic structure of the N-S cluster process and assuming that the multivariate PGF of the random vector:

$$[N(0, t_1], N(t_1, t_2], \dots, N(t_{k-1}, t_k)]$$

is desired, where $N(\cdot)$ counts the number of rainfall occurrences in the indicated nonoverlapping intervals, select

$$\xi(x) = 1 - \sum_{i=1}^k (1 - z_i) I_{(t_{i-1}, t_i]}(x) \quad (7)$$

which, after substituting in (2), yields

$$\begin{aligned} g(z_1, \dots, z_k) \\ = \exp \left\{ \mu \int_{-\infty}^{t_k} \left[g_{N_i} \left(1 - \sum_{i=1}^k (1 - z_i) p_i(t) \right) - 1 \right] dt \right\} \end{aligned} \quad (8)$$

where:

$$p_i(t) = \int_{t_{i-1}}^{t_i} f_T(\zeta - t) d\zeta \quad (9)$$

Equation (8) represents the multivariate PGF for the above random vector when the random variable $N(\cdot)$ is distributed according to the N-S cluster model; $p(t)$ and $p_i(t)$ represent the probability that a rainfall event, whose RGM occurred at time

t , falls within the time intervals indicated by the upper and lower limits of the integral. From this definition, it is clear that the nature of the pdf, $f_T(\zeta)$, determines the memory of the rainfall process.

The probability of any event can be obtained from (5) and (8). However, in this case, and as opposed to the Poisson process in which a complete stochastic description is obtained at the expense of the independence assumption, the N-S cluster model renders a general cluster dependence structure. In fact, the model expressed in (2), (5), and (8) has the Poisson process as one of its particular cases.

In this work, as by Kavvas and Delleur [1981], it is assumed that the random variable characterizing the number of rainfalls in a given cluster N_i follows a geometric distribution with parameter p . It is well known that the PGF of the geometric distribution is [Parzen, 1962]

$$g_{N_i}(z) = \frac{pz}{1 - (1 - p)z} \quad |z| < 1 \quad (10)$$

The distribution of the time positions of the individual rainfalls within their respective clusters, Y , is assumed to be negative exponential with parameter α , so that

$$\begin{aligned} f_T(\zeta - t) &= \alpha \exp[-\alpha(\zeta - t)] & \zeta - t > 0 \\ f_T(\zeta - t) &= 0 & \text{otherwise} \end{aligned} \quad (11)$$

The assumption of an exponential distribution for Y implies that no matter how far in the future from the time of occurrence of an RGM, there always exists a positive probability that a rainfall occurring at that time has been generated by the RGM which occurred in the infinite past.

Once the pdf's for N_i and Y have been chosen as the geometric distribution and the exponential distribution, the following expressions for the variance time function and the spectrum of counts are obtained:

$$\text{Var}[N(0, t_1)] = \frac{\mu(2 - p)t_1}{p^2} + \frac{2\mu(1 - p)}{\alpha p^2} (e^{-\alpha t_1} - 1) \quad (12)$$

and

$$g_+(w) = \frac{\mu}{\pi p} \left[1 + \frac{2(1 - p)}{p} \frac{\alpha^2}{\alpha^2 + w^2} \right] \quad w > 0 \quad (13)$$

The spectrum of counts is the Fourier transform of the covariance density,

$$C(u) = \lim_{\Delta t \rightarrow 0} \text{Cov} \frac{[\Delta N_t, \Delta N_{t+u}]}{\Delta t^2}$$

where $\Delta N_t = N(0, t + \Delta t] - N(0, t]$.

It is possible to obtain general expressions (distribution independent) for the above properties and for other moments like the mean $E[N(0, t_1)] = \mu E[N_i(\cdot)] t_1$, the rate of rainfall occurrence $\lambda(t_1) = \mu E[N_i(\cdot)]$, and the covariance function. The reader is referred to Ramirez and Bras [1982] or Vere-Jones [1970] for details.

FORECASTING AND CONDITIONAL PROBABILITY DISTRIBUTIONS OF THE N-S MODEL

Just as a point process can be described in two different ways, either in terms of its counting properties or in terms of its interval properties, forecasting problems for point processes are also of two different kinds. The first kind corresponds to the problem of finding the distributions of the number of events in future time intervals. The second kind corresponds

to the problem of finding the distributions of the time to and between various configurations of events. It is this last kind of problem that is dealt with in this work. The function of interest is the conditional distribution function (CDF) of the time to the next rainfall event, given some or all of the past of the process. The immediate past, in particular the time which has elapsed since the preceding rainfall event, is of utmost importance.

The following assumptions are in order: (1) the process is stationary, (2) the process has single occurrences, (3) the process has finite rate of occurrence, and (4) the complete stochastic structure of the process is known. All four assumptions hold true in the case of the N-S cluster model. Before proceeding, denote the finite dimensional distributions of the point process $N(\cdot)$ as

$$P_{i,j,\dots,k}(t_1, t_2, \dots, t_{n-1}, t_n) = \Pr[N(0, t_1] = i, N(t_1, t_2] = j, \dots, N(t_{n-1}, t_n] = k] \quad (14)$$

which are known from the pgfl of the N-S model.

Define X as the time to the next occurring rainfall event which is the quantity being forecasted. Its distribution shall be characterized in terms of the survivor function, defined as one minus the corresponding distribution function. Finally, the information on which the survivor function depends is taken in this work as the time elapsed since the occurrence of the last rainfall event. With this in mind, denote the survivor function as

$$S(x; H) = \Pr[X > x | H] = 1 - \Pr[X \leq x | H] \quad (15)$$

which gives the probability that the time X to the next rainfall event is greater than x conditional on the history H of the process.

In order to arrive at the desired CDF's, consider the worst situation or simplest case, namely, the situation when no information about the past of the process is available. According to this, (15) reduces to the unconditional probability:

$$S(x) = \Pr[X > x] \quad (16)$$

Using the duality between the counting properties and the interval properties of a point process, (16) can be rewritten as

$$S(x) = \Pr[X > x] = \Pr[N(0, x] = 0] = P_0(x) \quad (17)$$

Now, consider the case when the available information about the past of the process is that a period of time ζ has elapsed without any rainfall events occurring. For this case, (15) becomes

$$\Pr[X > x | N(0, \zeta] = 0] = \frac{\Pr[N(0, \zeta] = 0, N(\zeta, \zeta + x] = 0]}{\Pr[N(0, \zeta] = 0]} \quad (18)$$

where use of the fact $P(A|B) = P(A \cap B) / P(B)$ has been made. Equation (18) can also be written as

$$\Pr[X > x | N(0, \zeta] = 0] = \frac{\Pr[N(0, \zeta + x] = 0]}{\Pr[N(0, \zeta] = 0]} \quad (19)$$

Substituting (17) in (19) yields

$$S_0(x; \zeta) = \Pr[X > x | N(0, \zeta] = 0] = \frac{S(x + \zeta)}{S(\zeta)} = \frac{P_0(x + \zeta)}{P_0(\zeta)} \quad (20)$$

Equation (20) yields the CDF of the time to the next rainfall event, conditional on having observed the process for a period

of time ζ , without any rainfall events occurring. The subscript zero indicates that no events were observed during the interval ζ .

When the time of past rainfall occurrences is known, the situation complicates, since by the assumption of point occurrences the survivor function would now be conditioned on probabilities over infinitesimal intervals. To solve this problem, it is necessary to obtain the multivariate distributions of the time intervals between successive events [Jowett and Vere-Jones, 1972]. The theory of Palm functions asserts that a well-defined set of multivariate distributions for the intervals exists if the assumptions mentioned previously hold.

The Palm functions are defined as the following limits:

$$q_k(y) = \lim_{x \rightarrow 0} \Pr\{N(0, y] = k | N(-x, 0] > 0\} \quad (21)$$

$y > 0 \quad k \in Z_+$

When the intensity of the process defined as

$$\lambda = \lim_{x \rightarrow 0} \Pr\{N(0, x] > 0\} / x \quad (22)$$

is finite, these limits are shown to exist for stationary probability in the works of Khinchin [1955] and Daley and Vere-Jones [1970]. Observe that (21) simply yields the limit of the probability that k events occur in an interval $(0, y]$, given that at least one event has occurred in an arbitrary, immediately preceding interval $(-x, 0]$, as the length of this interval goes to zero. Also, observe that (21) expresses a relationship between a function of an interval measure y and a function of a counting measure $N(\cdot)$.

One final result obtained by Khinchin allows the Palm functions of (21) to be obtained as derivatives of the probability distribution $\Pr$ of $N(\cdot)$. Khinchin [1960] proves that for a stationary point process with single occurrences and finite intensity λ :

$$\begin{aligned} \Pr\{N(0, x] \leq k\} &= 1 - \lambda \int_0^x q_k(u) du \\ &= \lambda \int_x^\infty q_k(u) du \quad x > 0 \quad k \in Z_+ \end{aligned} \quad (23)$$

Denoting $\Pr\{N(0, x] \leq k\} \equiv P_k(x)$, (23) can be expressed in differential form as

$$DP_k(x) = -\lambda q_k(x) \quad (24)$$

Generalization of (23) leads ultimately to the highly desirable one-to-one relation between the description of a stochastic point process in terms of its counting properties and that in terms of its interval properties [Daley and Vere-Jones, 1970].

Now assume that the available information about the past of the process is restricted to the time ζ since the occurrence of the last rainfall event (an actual rainfall occurrence was observed at the beginning of interval ζ); the desired CDF is

$$S_{10}(x; \zeta) = \Pr[N(\zeta, \zeta + x] = 0 | \zeta \text{ time units have elapsed since last event}] \quad (25)$$

It is clear that (25) can be rewritten as

$$S_{10}(x; \zeta) = \lim_{h \rightarrow 0} \frac{P_{10}(h, h + \zeta + x)}{P_{10}(h, h + \zeta)} \quad (26)$$

From the definition of Palm functions in (21) and for $k = 0$,

$$q_0(x) = \lim_{h \rightarrow 0} \Pr\{N(0, x] = 0 | N(-h, 0] > 0\} \quad x > 0 \quad (27)$$

Now, observe that (27) can be expressed as

$$q_0(x) = \lim_{h \rightarrow 0} \frac{Pr\{N(-h, 0] > 0, N(0, x] = 0\}}{Pr\{N(-h, 0] > 0\}} \quad (28)$$

but since $Pr\{N(-h, 0] > 0\} = 1 - Pr\{N(-h, 0] = 0\}$, (28) reduces to

$$q_0(x) = \lim_{h \rightarrow 0} \frac{P_{10}(h, x + h)}{1 - P_0(h)} \quad (29)$$

Using the Palm-Khinchin theory and (24), (29) becomes

$$q_0(x) = \lim_{h \rightarrow 0} \frac{P_{10}(h, x + h)}{1 - P_0(h)} = -\frac{1}{\lambda} \frac{dP_0(x)}{dx} \quad (30)$$

Finally, using (30) in (26), we obtain for the desired CDF,

$$S_{10}(x; \zeta) = \frac{q_0(x + \zeta)}{q_0(\zeta)} = \left[\frac{dP_0(x + \zeta)}{dx} / \frac{dP_0(\zeta)}{dx} \right] \quad (31)$$

All the desired CDF's have now been obtained in terms of the survivor functions $S_0(x; \zeta)$ and $S_{10}(x; \zeta)$ given in (20) and (31), respectively. By an extension of the development presented above, it is possible to define Palm-type functions of higher order in which more complex situations for the past of the process can be accounted for [see Jowett and Vere-Jones, 1972; Ramirez and Bras, 1982].

Explicit expressions for (20) and (31) are obtained for the particular case of the N-S cluster process in which the cluster center process is Poisson with parameter μ , the distribution of the cluster size is geometric with parameter p , and the distribution of the time positions about cluster centers is exponential with parameter α .

To do so, the following pdf is needed:

$$P_0(x) = Pr\{N(0, x] = 0\} = g(z)|_{z=0} \quad (32)$$

Substitution in (5) of the PGF for $N(\cdot)$ given in (10) and of the pdf of Y given in (11) and using (32) yields

$$P_0(x) = \left[\frac{p}{1 - qe^{-\alpha x}} \right]^{\mu/\alpha} e^{-\mu x} \quad (33)$$

where $q = 1 - p$.

Substituting (33) in (20) yields

$$S_0(x; \zeta) = \left[\frac{1 - qe^{-\alpha \zeta}}{1 - qe^{-\alpha(\zeta + x)}} \right]^{\mu/\alpha} e^{-\mu x} \quad (34)$$

$S_0(x; \zeta)$ is the conditional probability that the time to the next rainfall X is greater than x , conditioned on the fact that the process has evolved for a period of time ζ without any events occurring and that storm arrivals follow the N-S model developed.

When a rainfall event occurs at the beginning of the interval ζ , (31) yields

$$S_{10}(x; \zeta) = \frac{q_0(x + \zeta)}{q_0(\zeta)} = \frac{[dP_0(x + \zeta)]/dx}{[dP_0(\zeta)]/dx} \quad (35)$$

Differentiating (33),

$$\frac{dP_0(x)}{dx} = -\mu e^{-\mu x} p^{\mu/\alpha} [1 - qe^{-\alpha x}]^{-(1 + \mu/\alpha)} \quad (36)$$

Now, evaluating (36) at $(x + \zeta)$ and ζ and substituting in (35) produces

$$S_{10}(x; \zeta) = \left[\frac{1 - qe^{-\alpha \zeta}}{1 - qe^{-\alpha(x + \zeta)}} \right]^{(1 + \mu/\alpha)} e^{-\mu x} \quad (37)$$

$S_{10}(x; \zeta)$ is the conditional probability that the time to the next rainfall X is greater than x , conditioned on the fact that ζ time units have elapsed since the last rainfall occurrence and that storm arrivals follow the N-S cluster model proposed.

Observe that for both (34) and (37), as the time ζ increases (time without rain or time since last rainfall), the probabilities $S_0(x; \zeta)$ and $S_{10}(x; \zeta)$ approach those of a Poisson model with parameter μ , meaning that in fact, the further in the future an event is considered, the more it starts to look as independent of the present or past of the process.

IRRIGATION SCHEDULING PROBLEM

This work follows the general formulation presented by Córdova and Bras [1981] and Bras and Córdova [1981] to solve the irrigation scheduling problem in regions of deficit irrigation. The solution couples a stochastic description of storm arrivals, intensities and durations with potential evapotranspiration, a crop model, and a one-dimensional soil root zone model. The coupling provides the means to describe the random fluctuations of soil moisture which in turn are the necessary input to an optimal control algorithm based on stochastic dynamic programming. In their work, the above authors assumed that storm arrivals were independent Poisson events. In contrast to the formulation of Bras and Córdova [1981] this work utilizes the conditional distributions of storm arrivals $S_0(x; \zeta)$ and $S_{10}(x; \zeta)$ derived in the previous section. The use of this conditional information requires a modification of the scheduling problem formulation.

The scheduling problem uses three interconnected subsystems: climate, soil, and plant. Except for the changes in the model of storm arrivals already alluded, all these subsystems are the same used by Córdova and Bras [1981] and Bras and Córdova [1981]. The reader is kindly requested to use those references to determine the details of the deterministic soil water balance and plant models. Similarly, details on potential evapotranspiration and probabilistic distributions of storm intensities and durations assumed are given in the above references or by Ramirez and Bras [1982]. Following is the reformulated control algorithm. In general, the control algorithm is composed of the following elements: a time scale T , a state space S , a control space C , an output space Y , a state transition function F , a stochastic law of motion P , and an output function G .

Time Scale T

The time horizon of the irrigation scheduling problem is the length of the irrigation season. At the beginning of the season the farmer has a limited but known volume of irrigation water. Thus the objective is to manage optimally this limited amount of water, so that net profits at the end of the season are maximized. The time scale can be expressed as

$$T = \{k: k = 1, 2, \dots, N + 1\} \quad (38)$$

where k represents days (decision stages) and N is the number of days in the growing season.

State Space S

A state space representation of a given system requires that the state variable chosen contains all the necessary information, besides the inputs, to determine future states of the system. Let X_k be the state vector, such that

$$\begin{aligned} X_k &\in S & k &\in T \\ X_k &= (\theta_k, \zeta_k, r_k) \end{aligned} \quad (39)$$

where θ_k is the soil moisture content at the root zone at beginning of decision stage (day) k ; ζ_k is the number of days, at decision stage k , since the occurrence of the last rainfall; and r_k is the volume of irrigation water available at beginning of decision stage (day) k .

The element ζ_k of the vector state $\mathbf{X}_k$ represents the knowledge the decision maker has about the immediate history of the precipitation process. Owing to the dependence in the rainfall occurrences, this history contains potentially valuable information about the future of storm arrivals. This information is encoded in the conditional distributions $S_{10}(x; \zeta)$ and $S_0(x; \zeta)$, derived in the previous section, that are to be included in the general irrigation model.

Control Space C

At each decision stage (days, weeks) during the irrigation season, the farmer (decision maker) has to decide whether or not to irrigate and how much. It is by making irrigation decisions that the farmer can control the soil moisture states. Thus the control variable is the amount of irrigation water U_k applied at decision stage k , such that

$$U_k \in C \quad k \in T \quad 0 \leq U_k \leq r_k \quad \forall k \in T \quad (40)$$

Output Space Y

The objective of the decision maker (farmer) is to maximize net profits at the end of the season. Net profits are a function of crop yield. Thus the output variable chosen is the contribution of decision stage k to total actual crop yield Y_k .

State Transition Function F

The motion of the system, expressed in terms of the dynamics of the vector state $\mathbf{X}_k$, is given by the state transition function

$$\mathbf{X}_{k+1} = F(\mathbf{X}_k) \quad k \in T \quad (41)$$

The state transition function is a composite function of three different operators or dynamic equations, one for each element of the state vector $\mathbf{X}_k$.

First, the soil moisture depletion operator, which yields the soil moisture content at the root zone at the end of the current decision stage, θ_{k+1} is

$$\theta_{k+1} = g[\theta_k(U_k), \Delta t] + V_k[\theta_k(U_k), \Delta t] \quad (42)$$

where V_k is the volume of water infiltrated from storm events during decision stage k ; Δt is the duration of decision stage k , taken as 1 day; and θ_{k+1} is the soil moisture content at the end of the stage. In (42) it is assumed that an irrigation decision has been taken at the beginning of the interval. Consequently, θ_k in (42) is the soil moisture content after the respective irrigation application.

The first right-hand member of (42), $g[\theta_k(U_k), \Delta t]$ represents the deterministic part of the soil moisture state transition from decision stage to decision stage and describes the dynamics of the soil moisture depletion process. The second member of the right-hand side of the equation, $V_k[\theta_k(U_k), \Delta t]$, represents the stochastic volume of water infiltrated from storm events during time interval Δt . Consequently, θ_{k+1} is random. Its probabilistic description is presented later as part of the stochastic law of motion, P .

It is necessary to emphasize that infiltration, either from storm events or irrigation applications, is assumed to occur instantaneously.

Second, the elapsed time operator, which yields the value of the variable ζ_{k+1} at the beginning of the next decision stage as a function of its present value, ζ_k , and of a random disturbance, W_k , is as follows:

$$\begin{aligned} \zeta_{k+1} &= h(\zeta_k, W_k) = \zeta_k + 1 & W_k &= 0 & k &\in T \\ \zeta_{k+1} &= h(\zeta_k, W_k) = \zeta_k & W_k &= 1 & k &\in T \end{aligned} \quad (43)$$

where W_k is a binary random variable that is equal to one if it rains during decision stage k or equal to zero otherwise.

Finally, the irrigation water operator given as a simple mass balance equation to guarantee that the volume of water used for irrigation during each decision stage k is less than or equal to the available volume of irrigation water:

$$r_{k+1} = l(r_k, U_k) \quad k \in T$$

such that

$$\begin{aligned} r_{k+1} &= r_k - U_k \\ r_{k+1} &\geq 0 \end{aligned} \quad (44)$$

Stochastic Law of Motion P

The stochastic law of motion P is a family of conditional distribution functions $P_k(\mathbf{X}_{k+1}|\mathbf{X}_k, U_k)$, which for each decision stage k yield the conditional distribution function of the state vector $\mathbf{X}_{k+1}$, conditional on its present value and the irrigation decision. Only two elements of the state transition vector $\mathbf{X}_{k+1}$ are stochastic, namely, the soil moisture content θ_{k+1} and the time in days since the occurrence of the last rainfall ζ_{k+1} . The state transition function for the available irrigation water, $l(\cdot, \cdot)$, is totally deterministic. Consequently, the stochastic law of motion involves only the description of the joint stochastic behavior of θ_{k+1} and ζ_{k+1} . To do so, the probability distribution function (pdf) of the terminal soil moisture content, as well as the pdf of the time since the last rainfall event have to be defined.

Soil moisture content (SMC) at the root zone is discretized, in the state space, from saturation SMC, to permanent wilting point (PWP). The discretization follows that of *Bras and Córdova* [1981]. It is a variable interval discretization whose index ranges from one at saturation to m at PWP. The length of each interval is such that under unforced conditions (no infiltration inputs), it takes one day for the depletion processes to drive the SMC from the upper bound to the lower bound of each interval, and m days to drive it from saturation to PWP. According to this, the pdf of the terminal soil moisture content, for all soil moisture states, can be expressed as a transition matrix:

$$\Phi(k) = \{\phi_{ij}(k)\} \quad i, j = 1, \dots, m \quad (45)$$

where $\phi_{ij}(k)$ represents the probability that at the end of decision stage k (or beginning of decision stage $k+1$) the soil moisture content is in state j , given that at the beginning of the stage it is in state i . The generic element of matrix $\Phi(k)$ is defined as

$$\begin{aligned} \phi_{ij}(k) &= P_{ij|w_k=0}(k)Pr[W_k=0|\zeta_k] \\ &\quad + P_{ij|w_k=1}(k)Pr[W_k=1|\zeta_k] \end{aligned} \quad (46)$$

where $P_{ij|w_k=0}(k)$ is the probability that at the end of stage k , the soil moisture content is in state j , given that at the beginning of the stage it is in state i , and no precipitation occurs during the stage. $P_{ij|w_k=1}(k)$ is its analogue for the case when precipitation occurs during the stage. Finally, $Pr[W_k=0|\zeta_k]$

and $Pr[W_k = 1|\zeta_k]$ are the conditional probabilities of no rain and rain during the given stage, respectively.

If there is no rain during decision stage k , then the infiltrated volume $V_k(\theta_k, \Delta t)$ is zero in (42). Thus for Δt equal to 1 day and according to the above discretization,

$$\theta_{k+1}^j = g[\theta_k^i(U_k), \Delta t] \quad (47)$$

and

$$j = i + 1$$

Consequently,

$$\begin{aligned} P_{i|j|W_k=0}(k) &= Pr\{\theta_{k+1} \in [v\theta_{k+1}^j, l\theta_{k+1}^j] | \theta_k \\ &\in [v\theta_k^i, l\theta_k^i] \text{ and } W_k = 0\} = 1 \quad j = i + 1 \\ P_{i|j|W_k=0}(k) &= Pr\{\theta_{k+1} \in [v\theta_{k+1}^j, l\theta_{k+1}^j] | \theta_k \\ &\in [v\theta_k^i, l\theta_k^i] \text{ and } W_k = 0\} = 0 \quad \text{otherwise} \end{aligned} \quad (48)$$

where the subscripts U and L stand for upper and lower bound of the given soil moisture interval.

When precipitation occurs during the time interval Δt , (k , $k + 1$), (42) can be written as

$$\theta_{k+1}^j = g[\theta_k^i(U_k), \Delta t] + V_k[\theta_k^i(U_k), \Delta t] = \theta_k^{i+1} + V_k^{ij} \quad (49)$$

In (49), soil moisture is first depleted from θ_k^i to θ_k^{i+1} and then replenished up to the value θ_{k+1}^j . Thus V_k^{ij} represents the volume of infiltrated water (from rainfall) required to bring the SMC from θ_k^{i+1} to θ_{k+1}^j . The upper and lower bounds of the j th interval can be expressed as

$$\begin{aligned} v\theta_{k+1}^j &= \theta_k^{i+1} + vV_k^{ij} \\ l\theta_{k+1}^j &= \theta_k^{i+1} + lV_k^{ij} \end{aligned} \quad (50)$$

Consequently,

$$\begin{aligned} P_{i|j|W_k=1}(k) &= Pr\{\theta_{k+1} \in [v\theta_{k+1}^j, l\theta_{k+1}^j] | \theta_k^i, W_k = 1\} \\ &= Pr\{lV_k^{ij} \leq V_k^{ij} \leq vV_k^{ij} | \theta_k \in [v\theta_k^i, l\theta_k^i] \text{ and } W_k = 1\} \end{aligned} \quad (51)$$

Finally,

$$\begin{aligned} P_{i|j|W_k=1}(k) &= \frac{1}{\Delta\theta_k^i} \int_{l\theta_k^i}^{v\theta_k^i} Pr\{lV_k^{ij} \leq V_k^{ij} \leq vV_k^{ij} | \theta_k \\ &\in [v\theta_k^i, l\theta_k^i], W_k = 1\} d\theta_k^i \end{aligned} \quad (52)$$

where

$$\Delta\theta_k^i = v\theta_k^i - l\theta_k^i \quad (53)$$

Expression (52) can be evaluated using (49) and (50) to calculate vV_k^{ij} and lV_k^{ij} and using an expression for the probability density function of the random variable V_k^{ij} , conditional on θ_k^i derived by Córdova and Bras [1981]. This expression depends only on the mean storm intensity and duration in the region of interest and on soil physical properties.

In a previous section, the conditional distributions $S_0(x; \zeta)$ and $S_{10}(x; \zeta)$ have been defined for a point process with single occurrences and applied to model storm arrivals. Assume that the time interval chosen, Δt equal to 1 day, between decision stages is sufficiently small to guarantee that

$$Pr[N(\Delta t) > 1] \approx 0 \quad (54)$$

where $N(\Delta t)$ counts the number of rainfall occurrences in interval Δt .

By definition,

$$\begin{aligned} S_0(x; \zeta) &= Pr[X > x | \zeta] \\ S_{10}(x; \zeta) &= Pr[X > x | \zeta] \end{aligned} \quad (55)$$

where X represents the time to the next rainfall occurrence.

Also, because of the duality between the definition of a point process as a counting process or as an interval process, (55) can be written as

$$Pr[X > x] = Pr[N(x) = 0] = 1 - \sum_{n=1}^{\infty} Pr[N(x) = n] \quad (56)$$

When x equals 1 day, (54) holds. Consequently, substituting in (56) and (55):

$$\begin{aligned} S_0(x = 1 \text{ day}; \zeta) &= Pr[N(x) = 0 | \zeta] = 1 - Pr[N(x) = 1 | \zeta] \\ S_{10}(x = 1 \text{ day}; \zeta) &= Pr[N(x) = 0 | \zeta] = 1 - Pr[N(x) = 1 | \zeta] \end{aligned} \quad (57)$$

Thus, the pdf of the random variable W_k can be obtained from (57) as

$$\begin{aligned} Pr[W_k = 0 | \zeta_k] &= S_0(x = 1 \text{ day}; \zeta_k) \\ Pr[W_k = 1 | \zeta_k] &= S_{10}(x = 1 \text{ day}; \zeta_k) \end{aligned} \quad (58)$$

and

$$\begin{aligned} Pr[W_k = 1 | \zeta_k] &= 1 - S_0(x = 1 \text{ day}; \zeta_k) \\ Pr[W_k = 1 | \zeta_k] &= 1 - S_{10}(x = 1 \text{ day}; \zeta_k) \end{aligned} \quad (59)$$

Expressions to evaluate $S_0(x; \zeta)$ and $S_{10}(x; \zeta)$ were given in the previous section.

This completes the definition of the SMC transition matrices. It should be noted that now, as opposed to Bras and Córdova [1981], the transition matrices $\Phi(k)$ are conditional on the random variable ζ_k . Also, it should be observed that the conditional information contained in the history of the precipitation process has now been effectively accounted for.

To define completely the stochastic law of motion, it is necessary to define the pdf of the random variable ζ_{k+1} , conditional on the value of ζ_k . It has already been stated that ζ_k represents time in days since the occurrence of the last rainfall. However, an upper bound on the value of ζ_k has not been defined. In order to do so, it is assumed that there exists a time lag in days beyond which the dependence in the rainfall arrival process is sufficiently weak to guarantee that

$$S_0(x; H) \neq S_0(x; H^*) = S_0(x; H_1) \quad 0 \leq H < H^* < H_1 \quad (60)$$

$$S_{10}(x; H) \neq S_{10}(x; H^*) \approx S_{10}(x; H_1) \quad 0 \leq H < H^* < H_1$$

In (60), H , H^* , and H_1 represent the knowledge the decision maker has about the immediate history of the storm arrival process. Thus the previous assumption simply states that there exists an upper bound H^* to the conditional information contained in the past of the process. According to this assumption, (43) becomes

$$\begin{aligned} \zeta_{k+1} &= h(\zeta_k, W_k) = \min(\zeta_k + 1, H^*) \quad W_k = 0 \quad k \in T \\ \zeta_{k+1} &= h(\zeta_k, W_k) = W_k \quad W_k = 1 \quad k \in T \end{aligned} \quad (61a)$$

and

$$\zeta_k \in \zeta_k^j \quad \text{if } \zeta_k = j \quad k \in T \quad j = 1, 2, \dots, H^* \quad (61b)$$

The conditional pdf of the random variable ζ_{k+1} , conditional on the value of ζ_k , can be expressed as a transition matrix:

$$\Psi(k) = \{\psi_{ij}(k)\} \quad i, j = 1, 2, \dots, H^* \quad (62)$$

where $\psi_{ij}(k)$ is the probability that at decision stage $k + 1$, ζ_{k+1} is in state j , given that at stage k , ζ_k is in state i .

The generic element $\psi_{ij}(k)$ can be expressed as

$$\psi_{ij}(k) = q_{ij|W_k=0}(k) Pr[W_k = 0|\zeta_k] + q_{ij|W_k=1}(k) Pr[W_k = 1|\zeta_k] \quad (63)$$

where

$$q_{ij|W_k=0}(k) = Pr[\zeta_{k+1} \in \zeta_{k+1}^j | \zeta_k \in \zeta_k^i \text{ and } W_k = 0] \quad (64)$$

and

$$q_{ij|W_k=1}(k) = Pr[\zeta_{k+1} \in \zeta_{k+1}^j | \zeta_k \in \zeta_k^i \text{ and } W_k = 1] \quad (65)$$

Expressions (64) and (65) represent the one-stage transition probabilities of the random variable ζ_k , conditional on whether or not it rains during decision stage k . These probabilities are evaluated using (61a) and (61b). They are

$$\begin{aligned} q_{ij|W_k=0}(k) &= 1 & j = i + 1 \\ q_{ij|W_k=0}(k) &= 0 & \text{otherwise} \end{aligned} \quad (66)$$

and

$$\begin{aligned} q_{ij|W_k=1}(k) &= 1 & j = 1 \\ q_{ij|W_k=1}(k) &= 0 & \text{otherwise} \end{aligned} \quad (67)$$

The joint stochastic behavior of the random variables θ and ζ can be obtained in terms of the transition matrices $\Phi(k)$ and $\Psi(k)$. If the SMC is in state i , θ_k^i , and the elapsed time is in state p , ζ_k^p ; then the probability that at the beginning of state $k + 1$, θ_{k+1} is in state j and ζ_{k+1} is in state q can be written as

$$Pr\{[\theta_k^i, \zeta_k^p] \rightarrow [\theta_{k+1}^j, \zeta_{k+1}^q]\} = \phi_{ij}(k) \psi_{pq}(k) \quad (68)$$

$$i, j = 1, \dots, m \quad p, q = 1, \dots, H^*$$

where $Pr\{[\cdot] \rightarrow [\cdot]\}$ stands for transition probability.

Generalizing (68) for all feasible values of the indexes i, j, p, q yields the desired stochastic law of motion.

Output Function G

The contribution of decision stage k to actual total crop yield has been defined as the output variable. The output function can be defined as

$$Y_k = G_k(X_k, X_{k+1}, U_k) \quad (69)$$

where Y_k is the contribution of decision stage k to total crop yield. G_k is the crop model presented by *Bras and Córdova* [1981]. Also see *Ramírez and Bras* [1982] for further detail.

Stochastic Dynamic Programming Solution

The objective is to maximize the expected value of the total net profits at the end of the growing season:

$$B^* = \max E \left[\sum_{k=1}^{N+1} R_k(X_k, X_{k+1}, U_k) \right] - PC \quad (70)$$

and

$$R_k(X_k, X_{k+1}, U_k) = PY_k - \beta U_k - \gamma C_k(U_k) \quad (71)$$

where P is the unit price of crop yield, Y_k is the contribution of decision stage k to total yield, β is unit cost of irrigation water, γ is the fixed cost of irrigation (labor cost), and U_k is the volume of irrigation water applied at decision stage k , PC are production costs different from irrigation costs, and

$$\begin{aligned} C_k(U_k) &= 1 & U_k > 0 \\ C_k(U_k) &= 0 & U_k = 0 \end{aligned}$$

When there is limited amount of water for irrigation, the stochastic dynamic programming (SDP) algorithm proceeds as follows:

Define

$$J_{N+1}^*[X_{N+1}] = J_{N+1}^*[\theta_{N+1}^i, \zeta_{N+1}^p, r_{N+1}] = 0 \quad (72)$$

$$\forall i \quad i = 1, \dots, m \quad \forall p \quad p = 1, \dots, H^*$$

By induction

$$\begin{aligned} J_k^*[\theta_k^i, \zeta_k^p, r_k] &= J_k[X_k, U_k^*] = \max_{U_k \in Q_k} \left\{ E[R_k(\theta_k^i, \zeta_k^p, U_k)] \right. \\ &\quad \left. + \sum_j \sum_q \phi_{ij}(k) \psi_{pq}(k) J_{k+1}^*[\theta_{k+1}^j, \zeta_{k+1}^q, r_{k+1}] \right\} \end{aligned} \quad (73)$$

and

$$r_{k+1} \leq r_k - U_k \quad (74)$$

Then, set

$$\Gamma_k^*(\theta_k^i, \zeta_k^p, r_k) = U_k^* \quad (75)$$

where Q_k is the set of feasible irrigation decisions at stage k and U_k^* is the optimal irrigation decision at stage k .

Obtain

$$J_1^*[\theta_1^i, \zeta_1^p, r_1] = B^*[\theta_1^i, \zeta_1^p, r_1] + PC \quad (76)$$

The SDP algorithm yields not only the expected maximum net benefits $J_k^*[X_k]$ for each decision stage and state vector but also, and most importantly, optimal decision functions $\Gamma_k^*(X_k)$ that give the optimal amount of irrigation water to be applied as a function of the SMC at the root zone θ_k , the time in days since the last rainfall occurrence ζ_k , and the volume of water available for irrigation r_k . When the available irrigation water r_1 is unlimited, there is no need to impose the mass balance constraint of (74) nor to carry r_k as part of the state vector X_k . Otherwise, the solution is the same.

MODEL CALIBRATION

Sources of Data

The parameters required by the soil-plant model were obtained from field experiments conducted at Colorado State University, with an early corn variety (Northrup King PX 20) in a uniformly deep Nunn clay loam soil during the growing season of 1974. This is the same case study used by *Córdova and Bras* [1981] and *Bras and Córdova* [1981]. The reader should go to those references for details on parameters of the soil and plant systems. *Ramírez and Bras* [1982] also summarize necessary data, with special emphasis on the work reported here.

Precipitation data were obtained from historical records at Denver, Colorado (NOAA station 05-2220). Economic coefficients were obtained by *Blank* [1975] from the economic study of the Fort Morgan irrigation area in Colorado and performed by *Conklin* [1974]. The growing season starts on May 15 and ends on September 11.

Climate Subsystem Model Calibration

Precipitation is modeled in terms of the process of storm arrivals and the storm intensity and duration. These parameters were estimated from 27 years of precipitation data at Denver, Colorado (1949–1975).

TABLE 1. Homogeneous precipitation periods

Precipitation Period	
I	May 15 to June 16
II	June 16 to July 18
III	July 18 to August 3
IV	August 3 to August 19
V	August 19 to September 11

In the work of *Córdova and Bras* [1981] and in order to justify the use of the Poisson process for the storm arrivals an independent storm event had to be defined. This was done by determining a minimum interstorm time that would yield a coefficient of variation equal to one in the implied one parameter exponential distribution. Neglecting seasonality in the precipitation process throughout the year, *Restrepo-Posada and Eagleson* [1982] determined that the minimum interstorm time was 17.7 hours for Denver, Colorado.

In this work, the objective is not to define a storm event that fits the assumptions of a prespecified model; rather, a logical storm event is defined, and then a precipitation model is fitted to the resulting time series. An interval size of 24 hours (1 day) is then used to define a rainfall occurrence. A storm event occurs in the middle of the day whenever the total precipitation on the given day is greater than 0.254 mm (0.01 inches).

The entire growing season (120 days) was divided into 15 consecutive 8-day-long periods to carry out the rainfall data analysis with respect to intensity and duration. First- and second-order moments for intensity and duration, as well as the number of storm events for each 8-day period, were obtained.

Comparing the number of storms in each of the 15 periods as well as the mean and variance of the storm intensity and duration, five arbitrary statistically homogeneous precipitation periods were defined. These homogeneous precipitation periods are as shown in Table 1. This means that although as stated in previous sections, the precipitation process is considered homogeneous (stationary) with respect to the storm arrivals, it is considered inhomogeneous with respect to the storm characteristics intensity and duration. Using the method of moments, an exponential distribution was fitted to storm intensity and storm duration in each homogeneous precipitation period. Also, for the sake of comparison, a homogeneous Poisson model was fitted to the process of storm arrivals. The parameters obtained for the analyzed data are as shown in Table 2.

Rainfall Arrivals Model Parameters

Three parameters are needed when the rainfall arrivals are modeled as a Neyman-Scott cluster process of the kind considered here. These are the parameter μ of the Poisson distribution for the parent process or rate of occurrence of RGM's.

TABLE 2. Parameters of Exponential Distribution for Storm Intensity, Duration, and Poisson Arrival Rate

	Homogeneous Precipitation Periods				
	I	II	III	IV	V
Storm intensity parameter	0.78	0.65	0.48	0.65	0.70
Storm duration parameter	0.24	0.43	0.36	0.46	0.37
Storm arrival rate	0.36	0.26	0.34	0.28	0.23

Poisson model parameter is 0.294.

TABLE 3. Fit to Log-Survivor Function and Spectrum of Counts

Parameter	Value
α	0.755296
p	0.674060
μ	0.234509

the parameter p of the geometric distribution for the cluster sizes, and the parameter α of the exponential distribution for the times of occurrence of the individual storms within their respective clusters. The calibration of these three parameters is done by fitting the theoretical spectrum of counts and log-survivor function for the N-S model to their estimated counterparts. The estimation of the spectrum of counts and of the log-survivor function is performed using the theory developed by *Cox and Lewis* [1966] for the statistical analysis of series of events and the computer program SASE by *Lewis et al.* [1969].

Estimation of the mean arrival rates throughout the year for the available data indicates a strong, annual cycle and possibly some higher-frequency variations. A single N-S model would not then be applicable for all seasons. Therefore the N-S model is used only on the 4-month irrigation season, assuming it is a stationary period with constant arrival rate.

In order to preserve the dependence structure, the model is fitted to the estimated spectrum of counts, and in order to preserve the structure of the interarrival times and the clustering behavior, it is fitted to the estimated log-survivor function.

The theoretical normalized spectrum of counts can be expressed, from (13), as

$$\frac{\pi p g_+(w)}{\mu} = 1 + \frac{2(1-p)}{p} \frac{\alpha^2}{\alpha^2 + w^2} \quad w > 0 \quad (77)$$

The theoretical unconditional log survivor function can easily be obtained from (37) by setting $\zeta = 0$ and taking natural logarithms. Thus the log-survivor function is

$$\ln S_{10}(x; 0) = \ln \Pr[X > x] = -\mu x + \left(\frac{\mu}{\alpha} + 1\right) \ln \frac{p}{1 - qe^{-\alpha x}} \quad (78)$$

The fitting of both (77) and (78) was performed using a nonlinear adaptive least squares algorithm developed at the Sloan School of Management at The Massachusetts Institute of Technology [Dennis et al., 1979]. As can be inferred, the fitting has to be simultaneous, since both functions have common parameters. The chosen parameter set is as shown in Table 3.

Figure 1 shows the fit to the spectrum of counts. There is no question that the representation is good. Figure 1 also shows the normalized spectrum of counts and 99% confidence intervals for the homogeneous Poisson process when fitted to the data from Denver. The large number of low frequencies outside of these intervals seem to indicate that the Poisson independence assumption is not valid. The value of α determined implies a rapid decay of the exponential distribution of Y , indicating the anticipated short memory of the process. The fit to the log-survivor function obtained is shown in Figure 2. The fit is rather good for the longer interarrival times, especially when compared to the fit that would be obtained using the Poisson assumption which implies a straight line.

Computation of the variance-time function supports the ob-

Fig. 1. Calibrated normalized spectrum of counts.

served non-Poisson behavior. For a homogeneous Poisson process the variance-time function plots as a straight line with slope equal to the rate of arrival. Comparison between the estimated variance-time function and the theoretical variance-time function under the Poisson assumption is shown in Figure 3. It shows not only the deviation from the Poisson case of the estimated variance-time function but also indicates an overdispersion and a clustering of the rainfall events [Vere-Jones, 1970].

Figure 3 also shows the behavior of the theoretical variance-time curve when the estimated parameters of the N-S model are used. Observe that the fit is at least as good as that obtained under the Poisson assumption. In this case the estimated N-S model produces a more overdispersed process than the actual one. However, since this fitting criteria is failed by both models, the departure being at least of the same magnitude for both models, model selection is not possible. On the other hand, the Poisson model only reproduces the rate of arrival, while the fitted N-S model reproduces the spectrum of

counts and the log-survivor function in addition to the rate of arrival. Overall, the judgment favors the N-S over the Poisson process. However, this judgment is subjective.

Any model that pretends to make use of the conditional information contained in the immediate history of rainfall arrivals should have at least the characteristic of being able to identify local short-term trends in the process. For instance, if during any given period, several days have passed without any storm occurring, the probability of rain should increase. However, when the number of days without rain keeps increasing, the model should be able to interpret this information as an indication of a drought, thus the probability of future rain should decrease.

In the framework of the CDF's expressed in (34) and (37), the above characteristic can be interpreted as follows. As ζ increases, that is, as the number of days since the last rainfall event increases, the model should detect a possible drought and yield a decreasing probability of rain in the future. This behavior can be observed in Figure 4 by fixing the value of x

Fig. 2. Calibrated log-survivor function.

Fig. 3. Estimated variance-time function.

and varying ζ . Observe that as ζ increases, $S_{10}(x; \zeta)$ also increases. On the other hand, for a fixed ζ , as x increases, $S_{10}(x; \zeta)$ decreases, meaning that the probability of rain in the future increases as we look further and further ahead (Figure 4). Thus, the simple model devised is capable of adequately modeling and forecasting short-term trends in the process. However, its simplicity neglects a great deal of valuable information by only considering the time since the last rainfall event. In spite of this, due to the short memory of the process expressed by the parameter α , the influence of the neglected information can probably be considered negligible.

CASE STUDY RESULTS

For the limited water case, five feasible irrigation decisions were considered at each decision stage: irrigate up to field capacity (FC), to 3/4 FC, to 1/2 FC, to 1/4 FC, or not to irrigate. The growing season was divided in 120 decision stages (daily irrigation decisions). The solution is given as optimal decision functions that yield the optimal amount of water to be applied at each decision stage as a function of the SMC and the number of days elapsed since the last rainfall

occurrence. For the limited water supply case, 20 cases with different total irrigation water available during the growing season were considered. The amount of water available ranged from 475.0 to 0.0 mm, and as implied above, discrete intervals of 25 mm were used. In this case, the optimal decision is also a function of the available irrigation water. At the upper levels of available water (475.0 mm) the solution practically corresponds to that of the unlimited water case and becomes independent of r_k .

The effect of the conditional Neyman-Scott model, encoded in the CDF's $S_0(x; \zeta)$ and $S_{10}(x; \zeta)$, on irrigation control is compared to results from four other precipitation models. Namely, the unconditional N-S model encoded in $P_0(x)$, the homogeneous and nonhomogeneous Poisson models whose parameters are as presented in Tables 2 and 4, and Córdova and Bras' [1981] nonhomogeneous Poisson model whose parameters are as presented in Table 4. It is necessary to make clear that the latter model was fitted by Córdova and Bras using a storm definition differing from that used here so that the comparison is not valid. However, in extending Córdova and Bras' model from weekly to daily irrigation decisions results are obtained which contradict previously accepted facts, stated by Blank [1975] and Bras and Córdova [1981], about irrigation on fixed dates or variable dates. The inclusion of Córdova and Bras' model only serves the purpose of showing

Fig. 4. Calibrated conditional distribution function.

TABLE 4. Total Amount of Precipitation Under the N-S and Poisson Models

	Precipitation Period					
	I	II	III	IV	V	Mean
<i>Poisson Model*</i>						
Mean storm intensity, mm/h	1.28	1.54	2.27	1.52	1.47	1.62
Mean storm duration h/storm	6.67	2.33	3.13	2.33	3.70	3.63
Mean rate of arrival	0.262	0.228	0.308	0.266	0.184	0.250
<i>N-S Model</i>						
Mean storm intensity, mm/h	1.28	1.54	2.08	1.54	1.43	1.57
Mean storm duration, h/storm	4.16	2.32	2.78	2.17	2.70	2.83
Mean rate of arrival	0.356	0.263	0.345	0.282	0.225	0.294

*Obtained from Córdova and Bras [1981].

Fig. 5a. Limited water case: maximum net benefits.

the effect of the particular data definition used on the results of the control algorithm.

Figures 5-7 summarize the results obtained with the different models and different initial conditions. Observe that, in general, all models analyzed behaved very consistently. The unconditional form of the N-S model is compared first to the conditional N-S model, and then to the Poisson model. As expected, the conditional N-S model yields greater expected net benefits at the end of the growing season than the unconditional N-S model for all the cases considered (Figures 5b, 6b, and 7b). This is obviously the result of water conservation, obtained through improved irrigation water management made possible by accounting for the conditional information contained in the precipitation process. The net worth of the conditional information contained in the rainfall process is the difference between the expected net benefits obtained using the conditional and the unconditional N-S model. The net worth increases with decreasing initial soil moisture content and de-

creases to a lower bound with increasing available irrigation water.

However, the unconditional N-S model always yields lower expected net benefits than the homogeneous Poisson model. This result is to be expected. When there is no available irrigation water, the system is being driven solely by rainfall. The longer interarrival times, produced by the clustering encoded in the N-S process, coupled to the soil moisture depletion processes and no possible control, lead to longer and more severe periods of water stress on the crop, thus producing reduced maximum expected net benefits as compared to those obtained with the Poisson model. The same reasoning applies when control is possible. Greater stresses induce larger irrigation water use leading to lower expected net benefits. The difference in expected net benefits at the end of the growing season can be quite large (Figures 5b, 6b, and 7b). The expected opportunity loss of choosing the wrong model (represented by this difference) increases with decreasing initial soil

Fig. 5b. Limited water case: maximum net benefits.

Fig. 6a. Limited water case: maximum net benefits.

moisture content and decreases with increasing available irrigation water. The most crucial situation, thus, corresponds precisely to the situation where improved irrigation control is needed, namely, dry initial soil moisture conditions and little available irrigation water. The issue of actually determining the true underlying precipitation process (or the model that most closely represents it), which is not dealt with in this paper, becomes of paramount importance for optimal resource allocation and managerial decision making.

For this case study, the conditional N-S model produces the same expected net benefits as the homogeneous Poisson model. First, this indicates that if the true precipitation process were a N-S process then its conditional information, when optimally used through a control system like the one presented, is capable of offsetting the effect of clustering. This leads to expected net benefits similar to those for an unconditional but wetter precipitation process. Second, if the dependence structure of precipitation arrivals were stronger than that exhibited by this data set the conditional N-S model would lead to even greater net benefits (Figures 5b, 6b, and 7b). Observe that for the case study, for dry initial soil moisture conditions, the net

benefits of the conditional N-S model are higher than for the Poisson model.

The difference observed between the conditional N-S model and the nonhomogeneous Poisson model also raises once more the question of the implied stationarity of the former model. Although it has been shown that the N-S model is a better representation of the rainfall occurrences than the homogeneous Poisson model, this has not been done with respect to the nonhomogeneous Poisson model. The question of which is really the true underlying precipitation process can only be partially answered by simulating the system with the different optimal policies obtained from each model. An interesting hypothesis is that the low frequency dependence observed in Figures 1-2 may be due to hard to detect nonhomogeneities in storm arrivals and not to clustering, as argued here and in past literature.

Some additional results are now simply stated. Observe from Figures 5-7 that the nonhomogeneous Poisson model fitted by *Córdova and Bras* [1981] always produces greater expected net benefits than the nonhomogeneous Poisson models fitted in this work, although the data used in both fits

Fig. 6b. Limited water case: maximum net benefits.

Fig. 7a. Limited water case: maximum net benefits.

were the same. However, Córdoba and Bras transformed the data by defining an independent storm event to justify the Poisson assumption. Table 4 compares the parameters obtained in both fits. In general, their model expects more precipitation. Also observe from Figures 5-7 that contrary to facts stated by Blank [1975] and Bras and Córdoba [1981], there is an increase in the expected net benefits when going from irrigation on fixed dates (weekly irrigations) to irrigation on variable dates (daily irrigation decisions). This result emphasizes the major effect that prior data manipulation and transformations can have on chosen models and ultimate results.

Finally, it is necessary to emphasize that the results presented in this section are only theoretical results based on hypothesized rainfall models. The results do not say anything about the adequacy of the representation of rainfall arrivals, only about the implied benefits if the hypothesized model were correct. Although there is no doubt that the N-S model is a better representation of the rainfall, operationally it may be just as good to use a homogeneous Poisson for this data set.

The superior performance of a nonhomogeneous Poisson assumption just indicates that if such nonparsimonious models were indeed better and true representation of the rainfall arrivals, then its corresponding operating rules would be most efficient.

CONCLUSIONS

The use of the Palm-Khinchin theory allows the derivation of general expressions for the conditional distribution functions of the time to the next rainfall event, conditional on part of the immediate history of storm arrivals. Explicit forms of these expressions for a particular form of the N-S model were derived. These functions are seen to possess the required characteristics for adequate short-term trend forecasting, particularly drought trends.

The dynamic precipitation model devised, with the capability of changing as the growing season progresses and of using the conditional information to predict future behavior, was coupled with a model of the plant-soil system. This coupling allowed the implementation of a stochastic control model to

Fig. 7b. Limited water case: maximum net benefits.

obtain optimal irrigation decisions. It is shown that for the case study, the added complexity of the precipitation model apparently results in a substantial improvement in the representation of the process, although this does not lead to significant differences in the theoretical expected maximum net benefits when compared to the homogeneous Poisson model. This implies that the time constant of the soil-plant system overwhelms the persistence of the Denver rainfall data.

The results of this work do not allow any conclusion as to which model is a better representation of the underlying precipitation process. Under the N-S assumption it is shown that the conditional information contained in the history of precipitation arrivals has a net economic worth. This economic worth is represented in additional net benefits at the end of the growing season. This net worth is by no means negligible and will increase steadily with time as the demand for water and food increases.

However, it is also shown that the difference in optimal expected net benefits can be quite large depending on whether precipitation is modeled as a Poisson process or as a N-S process. Since for this data set both models appear to be at least as good as each other, the large expected opportunity loss of choosing the wrong model indicates that the crucial issue of actually determining the true underlying precipitation process is of paramount importance for resource allocation and decision making. Lower or higher expected net benefits only reflect what the chosen precipitation model implies about the nature of rainfall. In order to determine which model better represents the actual process, theoretical results as presented here must be compared to simulations of the irrigation rules obtained under very long historical rainfall inputs. However, even after simulation of the irrigation operations, the results of some models may not differ depending on the climatic persistence and time constants of the soil-climate-plant model. Further study of the apparent dependence that could be induced by nonhomogeneities in storm arrivals is also required.

With the use of the Palm-Khinchin theory, it has been shown how conditional distribution functions of the time to the next rainfall event could be derived. However, the situations considered here were simple. In fact, only the time since the last rainfall event is accounted for. It is desirable to include more complex situations, when more than one event in the past is considered. In this way, the emphasis is not so much on the drought conditions but also on the wet conditions. The derivation of CDF's accounting for more complex situations seems possible by defining Palm-type functions of higher order. Ramirez and Bras [1982] give an example. However, from the onset, the complexity of the expressions obtained, as well as the implementation of a control model able to use the added information, makes this a difficult task.

Acknowledgments. This work was supported by the National Science Foundation under NSF grants 78-20245 ENG and CEE8211469. Work was performed at the Ralph M. Parsons Laboratory, Department of Civil Engineering, Massachusetts Institute of Technology. Portions of this paper were written while Rafael Bras was on sabbatical leave, sponsored by John Simon Guggenheim Foundation.

REFERENCES

- Blank, H., Optimal irrigation decisions with limited water, Ph.D. dissertation, Colo. State Univ., Fort Collins, 1975.
- Bras, R. L., and J. R. Córdova, Intraseasonal water allocation in deficit irrigation, *Water Resour. Res.*, 17(4), 866-874, 1981.
- Conklin, L. R., Farm and irrigation management in the South Platte Valley of Colorado, Masters thesis, Colo. State Univ., Fort Collins, Aug. 1974.
- Córdova, J. R., and R. L. Bras, Physically based probabilistic models of infiltration, soil moisture, and actual evapotranspiration, *Water Resour. Res.*, 17(1), 93-106, 1981.
- Cox, D. R., and P. A. W. Lewis, *The Statistical Analysis of Series of Events*, Barnes and Noble, New York, 1966.
- Daley, D. J., and D. Vere-Jones, A survey of the mathematical theory of point processes, in *Stochastic Point Processes*, edited by P. A. W. Lewis, John Wiley, New York, 1970.
- Dennis, J. E., Jr., D. M. Gay, and R. E. Welsch, An adaptive nonlinear least squares algorithm, *Tech. Rep. 1*, Alfred P. Sloan Sch. of Manage., Mass. Inst. of Technol., Cambridge, 1979.
- Gabriel, K. R., and J. Newmann, On a distribution of weather cycles by lengths, *Q. J. R. Meteorol. Soc.*, 83, 375-380, 1957.
- Gabriel, K. R., and J. Newmann, A Markov chain model for daily rainfall occurrence at Tel Aviv, *Q. J. R. Meteorol. Soc.*, 88, 90-95, 1962.
- Gupta, V. K., and E. C. Waymire, A stochastic kinematic study of subsynoptic space-time rainfall, *Water Resour. Res.*, 15(3), 637-644, 1979.
- Jowett, J., and D. Vere-Jones, The prediction of point processes, in *Stochastic Point Processes*, edited by P. A. W. Lewis, John Wiley, New York, 1972.
- Kavvas, M. L., Stochastic trigger models for flood peaks, 1, Development of model, *Water Resour. Res.*, 18(2), 383-398, 1982a.
- Kavvas, M. L., Stochastic trigger models for flood peaks, 2, Application of the model to the flood peaks of Goksu Karahacili, *Water Resour. Res.*, 18(2), 399-411, 1982b.
- Kavvas, M. L., and J. W. Delleur, The stochastic and chronologic structure of rainfall sequences: Application to Indiana, *Tech. Rep. 57*, Water Resour. Res. Cent., Purdue Univ., West Lafayette, Indiana, 1975.
- Kavvas, M. L., and J. W. Delleur, A stochastic cluster model of daily rainfall sequences, *Water Resour. Res.*, 17(4), 1151-1160, 1981.
- Khinchin, A. Y., *Mathematical Methods in the Theory of Queueing*, English translation, Griffin, London, 1960.
- Le Cam, L., A stochastic description of precipitation, in *IV Berkeley Symposium on Mathematical Statistics and Probability Proceedings*, edited by J. Neyman, University of California Press, Berkeley, 1961.
- Lewis, P. A. W., A. M. Katcher, and A. H. Weis, SASE IV—An improved program for the analysis of series of events, *Rep. RC 2365*, IBM, Yorktown Heights, N. Y., 1969.
- Moyal, J. E., The general theory of stochastic population processes, *Acta Math.*, 108, 1-31, 1962.
- Neyman, J., and E. L. Scott, A theory of the spatial distribution of galaxies, *Astrophys. J.*, 116, 144-163, 1952.
- Neyman, J., and E. L. Scott, A statistical approach to problems of cosmology, *J. R. Stat. Soc. B*, 20, 1-43, 1958.
- Parzen, E., *Stochastic Processes*, Holden-Day, San Francisco, Calif., 1962.
- Ramirez, J. A., and R. L. Bras, Optimal irrigation control using stochastic cluster point processes for rainfall modeling and forecasting, *Tech. Rep. 275*, Ralph M. Parsons Lab., Dep. of Civ. Eng., Mass. Inst. of Technol., Cambridge, 1982.
- Restrepo-Posada, P. J., and P. S. Eagleson, Identification of independent rainstorms, *J. Hydrol.*, 55, 303-319, 1982.
- Smith, E. R., and H. A. Schreiber, Point processes of seasonal thunderstorm rainfall, 1, Distribution of rainfall events, *Water Resour. Res.*, 9(4), 841-884, 1973.
- Vere-Jones, D., Some applications of probability generating functionals to the study of input/output streams, *J. R. Stat. Soc. B*, 30, 321-333, 1968.
- Vere-Jones, D., Stochastic models for earthquake occurrences, *J. R. Stat. Soc. B*, 32, 1-62, 1970.
- Waymire, E., and V. K. Gupta, The mathematical structure of rainfall representations, 1, A review of the stochastic rainfall models, *Water Resour. Res.*, 17(5), 1261-1272, 1981a.
- Waymire, E., and V. K. Gupta, The mathematical structure of rainfall representations, 2, A review of the theory of point processes, *Water Resour. Res.*, 17(5), 1273-1285, 1981b.
- Waymire, E., and V. K. Gupta, The mathematical structure of rainfall representations, 3, Some applications of the point process theory to rainfall processes, *Water Resour. Res.*, 17(5), 1287-1294, 1981c.
- R. L. Bras, Department of Civil Engineering, Room 48-311, Massachusetts Institute of Technology, Cambridge, MA 02139.
- J. A. Ramirez, Department of Civil Engineering, Room 48-207, Massachusetts Institute of Technology, Cambridge, MA 02139.

(Received September 27, 1983;
revised November 30, 1984;
accepted December 4, 1984.)