ELSEVIER

Simulation Practice and Theory 3 (1996) 401–420

SIMULATION PRACTICE & THEORY

An integrated simulation model for construction

R.J. Wales [a], S.M. AbouRizk [b,*]

[a] North American Construction Group, 16745 – 111 Avenue, Edmonton, Alberta, Canada T5M 2S4
[b] Department of Civil Engineering, University of Alberta, Edmonton, Alberta, Canada T6G 2G7

Received 28 November 1994; revised 2 May 1995

Abstract

Systems simulation has proven to be an effective tool in the analysis of various manufacturing operations. Simulation has also found its way to construction as a means for project planning and control as well as for operation analysis. It was found especially useful for project managers in determining the impact of various factors upon the construction schedule. Weather, for example, influences the duration of many construction activities thereby causing uncertainty in the project schedule. This paper presents a simulation application where the authors have developed a combined simulation model that integrates many methods required for the successful application of simulation in modeling complex projects such as those encountered in construction. The methodology implemented includes using a general purpose simulation language where the user can extend the functionality of the system by writing his own code, developing the required methods for stochastically generating the occurrence of weather variables for use in project simulation studies, and estimating productivity using artificial intelligence.

Keywords: Simulation; Weather modeling; Construction; Risk analysis; Planning

1. Introduction

Construction projects, like many other projects undertaken in the open environment, are subject to effects of numerous factors leading to uncertainty in the timing and sequence of project activities. Uncertainty factors such as weather, soil conditions, resource availability, material arrivals, labor strikes, change orders, and others can be modeled as stochastic processes. It is advantageous from a management and planning perspective to model the occurrences of these factors, quantify their effects,

* Corresponding author.

SSDI 0928-4869(95)00018-6

and determine their impact on project schedule in order to improve the overall project plan.

Computer simulation is effective in performing these functions. Specifically, combined continuous-discrete event simulation modeling may be utilized, where, the network schedule (e.g. Critical Path Method (CPM), Project Evaluation and Review technique (PERT), or Precedence Diagramming Method (PDM)) is modeled in discrete event format, and other uncertainty factors as continuous stochastic processes. Project simulation modeling in this fashion requires:

(a) precise stochastic models capturing arrival processes of uncertainty factors,

(b) combined continuous-discrete event simulation modeling methods, and

(c) accurate quantification of uncertainty factors impact on productivity.

The objective of this paper is to introduce an integrated model that can enhance project scheduling and planning. The focus of this work will be on weather as it is representative of other continuous random phenomena encountered in construction. Weather is also the most involved of the uncertainty factors affecting 50% of construction activities [3]. Therefore, weather is sure to cause significant uncertainty in project scheduling. Similar modeling strategies may also be developed for other uncertainty factors (e.g. material delivery, soil conditions, etc.) following the same approach presented herein.

The modeling strategy employed takes advantage of available implementations so as not to duplicate already existing systems. The system is portable in that it only requires the SLAMSYSTEM program (available from Pritsker Corporation in commercial and academic versions) to run the simulation experiment. The process of converting PDM network schedules, solving for parameters of the stochastic weather generation model, and training of neural networks were automated and implemented for use on personal computers running Microsoft Windows. The programs are self contained requiring no other programs to run.

2. Background

Accurate forecast of project milestones and completion time allows management to establish the rate of completion of project activities and an accurate baseline to compare performance. A realistic forecast of project completion time is important since the schedule affects project costs. Variability in the project schedule makes estimating the project completion time a difficult task. This has led to techniques such as PERT and Monte-Carlo-based simulation which quantify the impacts of uncertainty variables. The PERT system assumes project activity durations are independent random variables (see [9], for example) making PERT probabilities unsuitable for many construction projects [5]. Many Monte-Carlo-based simulation models make the same assumption. Finally, managing uncertainty in the schedule is difficult since these methods do not establish the causes of uncertainty.

Carr presented a Monte-Carlo simulation based method that accounts for task interdependencies by maintaining time-dependent relationships among actual activity durations [5]. Ahuja and Nandukamar extended these concepts further and included uncertainty variable generation and quantification of their impact [1]. Both methods first simulate the impact of uncertainty variables independent of time of year and then simulate daily impacts of time-dependent variables. Improvements to these models can be achieved by implementing daily stochastic generation of uncertainty variables and quantification of uncertainty impacts through quantitative rather than qualitative methods.

This modeling implementation is applicable to many of the available general purpose simulation languages. SLAMSYSTEM [18] was chosen in this project for its combined discrete-event and continuous modeling capability.

3. Modeling method

Current project scheduling practices almost always include CPM, PDM or PERT schedules. In this model we take an existing network and automatically convert it to a SLAMSYSTEM simulation model in order to make use of current scheduling philosophies. Stochastic processes are then superimposed on the schedule to get a final accurate representation of the project schedule as shown in Fig. 1.

The model operates in the following sequence: weather conditions are generated for a given day; progress is estimated; and then credited to activities scheduled for that particular day. Progress is estimated by employing a trained neural network which accepts weather parameters as input and estimates productivity as output. When an activity is credited with enough progress it is complete and its followers may be scheduled for work. This process is detailed in the algorithm given in Fig. 2.

A sample of project completion times (or any milestone event) is obtained by repeating the experiment many times resulting is a sample that categorizes the project duration.

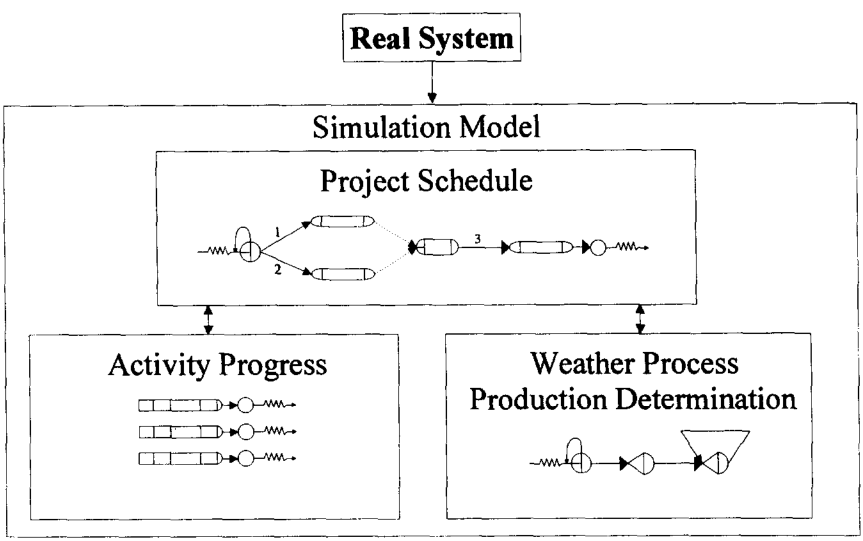

Fig. 1. Modeling concepts.

1. Start date
2. Begin simulation
3. Generate weather condition for the current day
4. Estimate progress based on weather conditions
5. Credit activity progress for activities scheduled for the current day
6. If activity work requirement is complete then go to step 7 otherwise advance clock by one day and go to step 3
7. Stop activity progress and allow follower activities to commence
8. If this is the final activity then end simulation, otherwise advance clock by one day and go to step 3

Fig. 2. Model algorithm.

The focus of this paper is the method for generating uncertainty variables. Combined continuous simulation modeling, as well as CPM to SLAMSYSTEM conversion are covered elsewhere (see for example [16]). In general terms, we make use of a number of systems to achieve integration as demonstrated in Fig. 3. Systems used for this purpose include the SLAMSYSTEM general purpose simulation language [15], the Visual Basic programming language for Windows, and neural network libraries, and Microsoft Project for Scheduling [13]. A project plan that is available in the form of a PDM network is converted into SLAMSYSTEM model using IS_CNV (Integrated System Converter) – a computer program written in Visual Basic. IS_FIT is another program that analyzes weather data and provides statistical distribution parameters for use in the SLAMSYSTEM simulation. Training for neural networks (NN) is also achieved in the self-contained program IS_NN which is also written in Visual Basic making use of the NeuroWindows libraries [14]. The network parameters determined in the program are then used in the SLAMSYSTEM simulation for estimating productivity achieved given certain weather conditions. The user needs to have access to SLAMSYSTEM to make use of the described model. Commercial, as well as academic versions of this product are readily available from Pritsker Corporation [18]. The programs are self-contained in executable packages requiring a personal computer running Microsoft

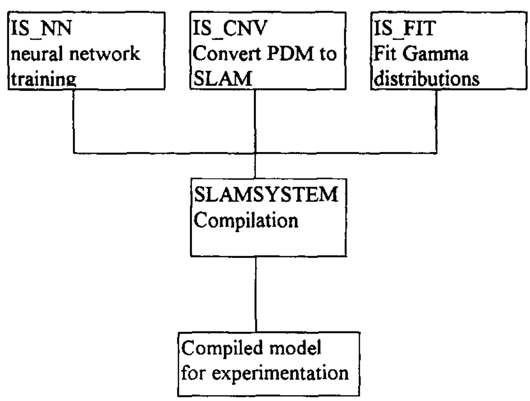

Fig. 3. Programs used in the modeling strategy and the analysis.

Windows 3.1. They are in the public domain and are available from the authors upon request.

4. Overview of the weather generation model

Weather variables of interest in construction simulation are those that affect productivity or may cause stoppage of work. This model generates precipitation, daily maximum temperature, and daily minimum temperature. Other variables (e.g., wind speed) can be generated by obtaining historical data records and following the same procedure. The weather generation scheme considered is a modification of the work described by Richardson [17]. Historical data is analyzed to determine the underlying stochastic process of the meteorological phenomena being modeled. The derived stochastic process is then used to generate weather variables for future dates. The basis of the model is precipitation which is considered as the primary variable. Other variables are adjusted according to precipitation and consequently to the wet or dry status of each day.

To model precipitation events a number of available procedures can be used. A Markov process that accomplishes this was described by Smith and Hancher [19]. Katz [10] and Kavvas [11] have also shown that first-order Markov chains are successful in these applications. Based on the above we selected the Markov process to model precipitation events in our model. In particular, a two-state (wet, dry) first-order Markov chain, as shown in Fig. 4 is used.

In order for the stochastic process to be useful it must be able to produce the status of the day, the minimum and maximum temperatures for the day, and in the event it is a wet day, the amount of precipitation for that day. A two-parameter Gamma distribution is fitted to historical records and then used in the simulation experiment to sample the appropriate amount of precipitation for the day. Seasonality of precipitation amounts is ensured by calculating separate parameters for the gamma distribution for each month of the year. During the project simulation experiment, the calendar date is maintained and therefore the appropriate parameters of the fitted Gamma distribution are used for sampling.

Generation of maximum and minimum temperatures utilizes the weekly stationary generating process suggested by Matalas [12]. This is a stochastic process which maintains dependencies between weather variables on the same day and from day to day.

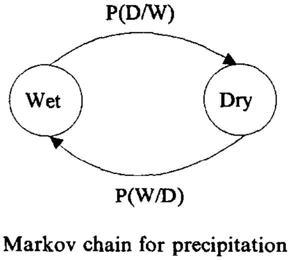

Fig. 4. Markov chain for precipitation states.

Generating daily weather variables during a simulation study follows the procedure depicted in Fig. 5. At the start of the simulation initial precipitation status is determined by generating a uniform random number on the interval 0 to 1. If the number generated is less than or equal to the probability of the day being wet, P(W), for the current month then day 1 of the simulation study is classified as wet (i.e. 0.2 mm or more precipitation) otherwise, the day is considered dry. In the event of a wet day, precipitation amount is determined by generating another uniform random number on the interval 0 to 1 and transforming to a gamma variate defined for the current month. Next, for the first day of a simulation study residuals from the

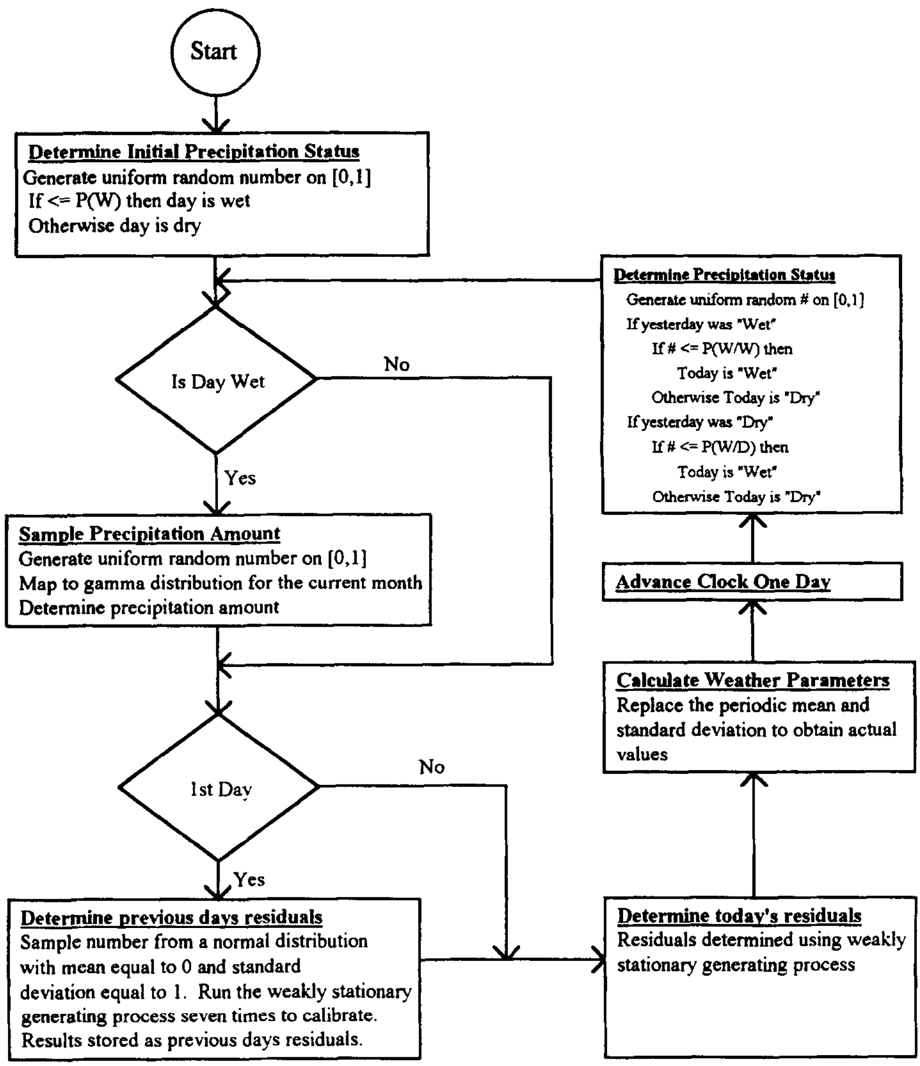

Fig. 5. Weather generation flowchart.

previous day must be determined. The residuals are sampled from a normal distribution with a mean of zero (0) and a variance of one (1). Today's residuals are determined from the weekly stationary generating process. From the generated residuals actual values are determined by multiplying the residual by the periodic standard deviation and adding the periodic mean. At this point, activity production is determined based on the generated parameters and activity progress is evaluated. The simulation clock is advanced by one day and precipitation status is determined for the current day using the transitional probabilities. This cycle is repeated until the project is completed.

5. Details of the stochastic weather generation model

Stochastic weather variable generation requires identifying properties of observed weather sequences. This allows the generation of numerous weather sequences for simulation experiments all having the same underlying properties. In this study, precipitation is considered the primary variable and all other variables are conditioned on the precipitation status of the day [17].

5.1. Precipitation component

Precipitation modeling uses a two-state (wet, dry) first-order Markov chain to determine the occurrence of wet or dry days and a two-parameter gamma distribution to determine precipitation amount. Precipitation status (wet or dry) of a given day is conditioned on the status of the previous day using a first-order Markov chain. Each day is classified as wet or dry depending on the amount of precipitation that has fallen. A day in which 0.2 millimeters or more precipitation has fallen is classified as a wet day. In the case of snowfall, equivalent water depth is used as a measure of precipitation amount.

The Markov chain process is modeled using the procedure presented by Smith and Hancher [19]. The two-state Markov chain is defined by transitional probabilities of moving from one state to another from one day to the next. Transitional probabilities are dependent on the time of year for most locations [17]. Therefore, separate transitional probabilities are calculated for each month of the year to reflect the seasonality of these parameters. Eq. (1) defines the transitional probabilities of transiting from state i on day d-1 to state j on day d.

$$P_{m,d}(j/i)=\frac{n_{ij}}{n_i} \tag{1}$$

The count of transitions from state i to state j is denoted by n_{ij} and n_i is the count of all transitions from state i to the other states for month m and x years of record.

When simulation begins the status of the previous day is not known. Therefore, the overall probability of a day being wet or dry must be determined in order to establish the initial precipitation status. The probability that any given day is in state i for month m can be calculated by Eq. (2).

$$P_m(i)=\frac{N_i}{X_m} \tag{2}$$

The count of all days observed in state i is denoted by N_i and X_m is the total number of days in x years of record for month m.

Clearly, it is only necessary to define two transitional probabilities and one overall probability to fully define the model. For this model $P_{m,d}(\mathrm{W/W})$, $P_{m,d}(\mathrm{W/D})$, and $P_m(\mathrm{W})$ are determined where: $P_{m,d}(\mathrm{W/W})$ is the probability of a wet day on day d given a wet day on day d-1 in month m; $P_{m,d}(\mathrm{W/D})$ is the probability of a wet day on day d given a dry day on day d-1 in month m; and $P_m(\mathrm{W})$ is the overall probability of a wet day in month m.

Generation of wet days requires determination of a precipitation amount. A two-parameter gamma distribution is fit to historical records through the moment matching procedure. Again, a separate gamma distribution is determined for each month of the year to reflect seasonality of precipitation amounts. The two-parameter gamma distribution is described by Eq. (3a).

$$f(x)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\exp\left[-\frac{x}{\beta}\right]x^{a-12}, \quad x>0, \tag{3a}$$

$$\mu=\alpha\beta \tag{3b}$$

$$\sigma^2=\alpha\beta^2 \tag{3c}$$

The mean (μ) and variance (σ^2) are related to the alpha (α) and beta (β) parameters by Eqs. (3b) and (3c). The mean and variance of the historical records is matched to that of the gamma distribution and the resulting alpha and beta parameters are calculated to describe the gamma distribution for each month.

Precipitation parameters are fully defined by calculating Markov chain transitional probabilities and gamma distribution parameters. Table 1 displays a summary of the precipitation variables for the Edmonton Municipal Airport Weather station based on the years 1972 to 1991, for illustration. Remaining weather variables are now determined based on the precipitation status of the day.

The precipitation component of the project simulation model operates according to the flowchart shown in Fig. 6. After simulation begins the precipitation status of the previous day must be determined in order to use the Markov transitional probabilities to determine the precipitation status for the current day. A random number is sampled from a uniform distribution between 0 and 1. If the random number is less than or equal to the probability of a wet day in the current month, $P_m(\mathrm{W})$, then the previous day is considered wet. Otherwise, the previous day is dry. Today's precipitation status is determined from the Markov transitional probabilities for the current month. Again, a random number is sampled from a uniform distribution between 0 and 1. If the previous day was wet, today is also wet if the random number is less than or equal to $P_m(\mathrm{W/W})$. Otherwise today is dry. If the previous

Table 1
Precipitation parameters

Month (1)	P_m (W/D) (2)	P_m (W/W) (3)	P_m (W) (4)	Alpha (α) (5)	Beta (β) (6)
January	0.433333	0.566667	0.353226	0.775529	2.234496
February	0.462857	0.537143	0.325000	1.214286	1.441091
March	0.596591	0.403409	0.287097	0.683679	2.701456
April	0.608696	0.391304	0.238333	0.710126	4.844945
May	0.493023	0.506977	0.358065	0.718201	6.102239
June	0.413534	0.586466	0.465000	0.662894	9.293379
July	0.479401	0.520599	0.440323	0.549877	12.664197
August	0.516129	0.483871	0.411290	0.681798	8.425219
September	0.477157	0.522843	0.340000	0.923529	5.314970
October	0.625000	0.375000	0.214516	0.794643	3.410225
November	0.534483	0.465517	0.298333	0.793266	1.771274
December	0.470046	0.529954	0.359677	0.728183	2.740692

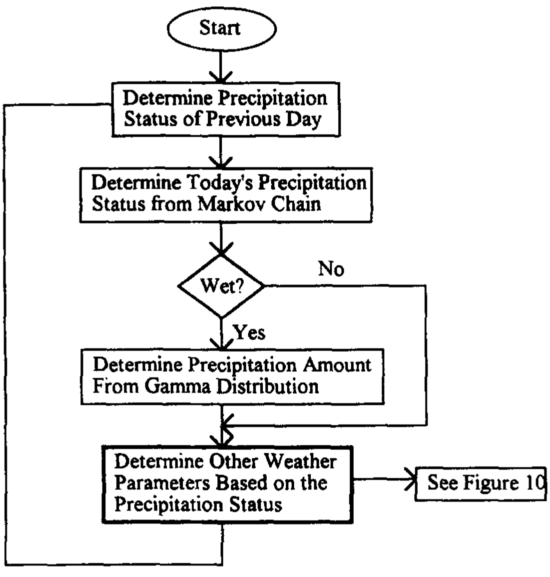

Fig. 6. Precipitation simulation procedure.

day was dry, today is wet if the random number is less than or equal to P_m(W/D). Otherwise the status of the current day is dry.

When wet days are generated a precipitation amount is determined from a two-parameter gamma distribution for the current month. Two uniform random numbers are generated between 0 and 1. The numbers are mapped onto the gamma function by a procedure that depends on the value of alpha (α) to determine a precipitation amount. For values of alpha greater than 0 and less than 1 the method of Jöhnk is employed [4]. When alpha is greater than or equal to 1 and less than 5 the method described in [6] as modified by Tadikamala [20] is employed. Although there were no cases of alpha being greater than or equal to 5 a weighted selection of Erlang samples would be employed [16]. Other weather variables are determined at this point according to the procedure described in the following section. Next, time advances one day and the precipitation status is determined from the Markov chain and the procedure repeats.

5.2. *Maximum and minimum temperature component*

The approach used in this component of the model considers maximum and minimum temperatures to be a continuous multivariate stochastic process with daily means and standard deviations conditioned on the precipitation status of the day [17]. Following the technique of Yevjevich [23] the historical time series records are first reduced to a series of residuals by removing the periodic mean and standard deviation. Eqs. (4a) and (4b) are used to determine residual elements of each series.

$$x_{y,d}(i)=\frac{X_{y,d}(i)-\bar{X}_d^0(i)}{\sigma_d^0(i)}, \quad P_{y,d}=0 \tag{4a}$$

$$x_{y,d}(i)=\frac{X_{y,d}(i)-\bar{X}_d^1(i)}{\sigma_d^1(i)}, \quad P_{y,d}>0 \tag{4b}$$

The subscripts y and d refer to the year and day respectively for each variable. The variable $X_{y,d}(i)$ is the residual element for variable i; $X_{y,d}(i)$ is the value of variable i; $\bar{X}_d^0(i)$ is the periodic mean of variable i on dry day d; $\sigma_d^1(i)$ is the periodic standard deviation of variable i on dry day d; $\bar{X}_d^1(i)$ is the periodic mean of variable i on wet day d; $\sigma_d^1(i)$ is the periodic standard deviation of variable i on wet day d; and $P_{y,d}$ is the precipitation in millimeters. Periodic means and standard deviations are determined by calculating the mean and standard deviation for each day of the year based on the wet or dry status of the day and smoothing the results with a general regression neural network. This results in a periodic mean curve for wet days, a periodic mean curve for dry days, a periodic standard deviation curve for wet days, and a periodic standard deviation curve for dry days. Periodic means and standard deviations in Eqs. (4a) and (4b) are determined from the smoothed curves. Fig. 7 demonstrates through a plot the mean daily maximum temperature for dry days for the Edmonton region as an illustration.

Residual elements are the basis of the weather generation scheme a plot illustration for the same data given in Fig. 7 is supplied in Fig. 8 for illustration as well. This work uses the weekly stationary generating procedure suggested by Matalas [12] which generates residual elements of the weather variables based on the residuals from the day before plus a random component. Actual values are determined by solving for in Eqs. (4a) and (4b) depending on the precipitation status. The weekly stationary generating process is defined by Eq. (5) for n weather variables.

$$x_{y,d-1}(i)=Ax_{y,d-1}(i)+B\varepsilon_{y,d}(i) \tag{5}$$

The variables $x_{y,d}(i)$ and $x_{y,d-1}(i)$ are $(n\times 1)$ matrices of residual elements for year

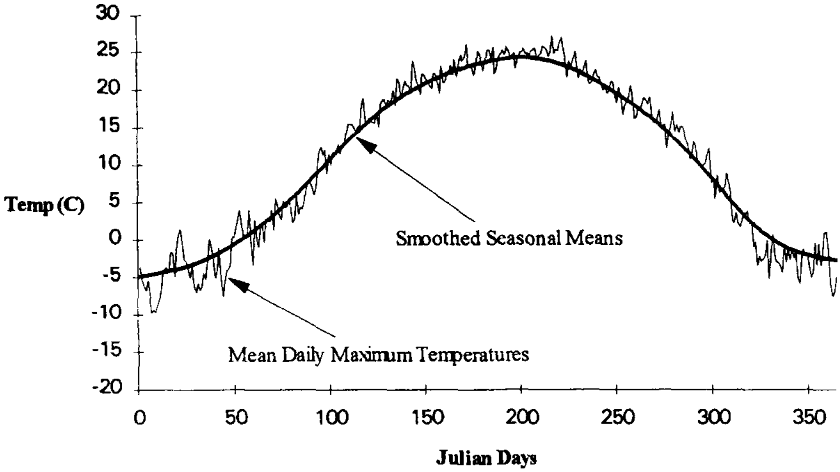

Fig. 7. Maximum temperature means (dry days).

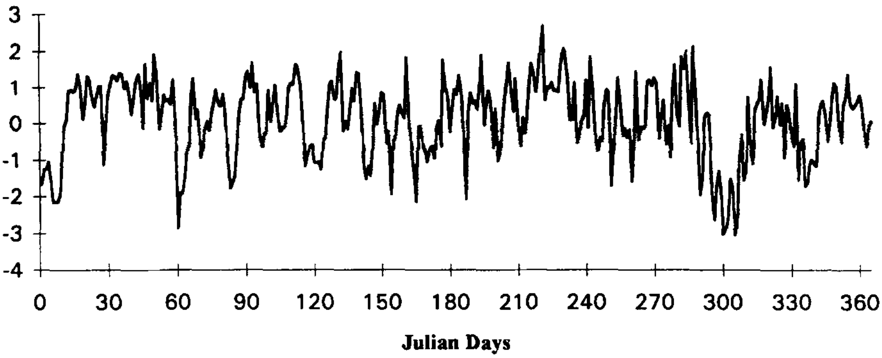

Fig. 8. Maximum temperature residual series (1991).

y and day d and $d-1$ for variables i to n; $\varepsilon_{y,d}(i)$ is a $(n \times 1)$ matrix of independent random components sampled from a normal distribution with a mean of 0 and a variance of 1; and A and B are $(n \times n)$ matrices defined such that the time dependence and interdependence structure of the historical residual series is preserved. Use of Eq. (5) implies that the residual series of each weather variable is normally distributed and a first-order linear autoregressive model describes the serial correlation of each variable [12].

A and B matrices ensure the generated residual series maintain the serial and cross-correlation characteristics of the historical series. Matrices A and B may be determined from Eqs. (6) and (7).

$$A = M_1 M_2^{-1}, \tag{6}$$

$$BB^{\mathrm{T}} = M_0 - M_1 M_0^{-1} M_1^{\mathrm{T}} \tag{7}$$

M_0 and M_1 are $(n \times n)$ lag 0 and lag 1 covariance matrices of the residual series. Variances of the residual series were approximately equal to 1, therefore M_0 and M_1 are essentially matrices containing lag 0 and lag 1 cross-correlation coefficients. M_0 and M_1 are defined by (8) and (9).

$$M_0 = \begin{bmatrix} 1 & \rho_0(1,2) & \dots & \rho_0(i,j) \\ \rho_0(2,1) & 1 & & \\ & & 1 & \\ \rho_0(j,i) & \dots & \dots & 1 \end{bmatrix} \tag{8}$$

$$M_1 = \begin{bmatrix} \rho_1(1) & \rho_1(1,2) & \dots & \rho_1(i,j) \\ \rho_1(2,1) & \rho_1(2) & & \\ & & & \\ \rho_1(j,i) & \dots & \dots & \rho_1(i) \end{bmatrix} \tag{9}$$

The variable $\rho_0(i,j)$ is the lag 0 cross correlation coefficient between variables i and j; $\rho_1(i)$ is the lag 1 serial correlation coefficient for variable i; $\rho_1(i,j)$ is the cross correlation coefficient between variables i and j with variable j lagged one day with respect to variable i. For this example only maximum and minimum temperature are being generated so M_0 and M_1 are 2×2 matrices. For the Edmonton Municipal Airport Weather Station these values were computed for maximum and minimum temperature series for the years 1972 to 1991 and were found to have the values shown below.

$$M_0 = \begin{bmatrix} 1.000000 & 0.962192 \\ 0.962192 & 1.000000 \end{bmatrix}, \quad M_1 = \begin{bmatrix} 0.626742 & 0.691539 \\ 0.619725 & 0.700515 \end{bmatrix}.$$

Solving for the values of the A matrix is simple and straightforward, however, solving for the B matrix values is more complicated due to the construction of the matrix equation. Following the solution procedure suggested by Young [24] B is assumed to be a lower triangular matrix. Matrix C is defined by Eq. (10). In the case of a 2×2 matrix Eq. (10) takes the form of the equation shown in (11). The elements of B can be determined from Eqs. (12), (13), and (14).

$$C = BB^{\mathrm{T}} = M_0 - M_1 M_0^{-1} M_1, \tag{10}$$

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \tag{11}$$

$$b_{11} = \sqrt{c_{11}}, \tag{12}$$

$$b_{21} = \frac{c_{21}}{b_{11}}, \tag{13}$$

$$b_{22} = \sqrt{c_{22} - b_{21}^2}. \tag{14}$$

For matrices having dimensions greater than 2×2 one can simply deduce the proper solutions for the elements of the B matrix using the same procedure. Using the equations above, the elements of A and B for the sample data were determined and are shown below.

$$A=\begin{bmatrix} -0.521001 & 1.192842 \\ -0.732002 & 1.404841 \end{bmatrix}, \quad B=\begin{bmatrix} 0.708263 & 0 \\ 0.634603 & 0.258470 \end{bmatrix}.$$

The elements of Eq. (5) are now fully defined. Generation of maximum and minimum temperatures follows the process shown in Fig. 9. The temperature routine is called from the precipitation routine of Fig. 6. The first step is to sample two random numbers from a normal distribution with a mean of zero (0) and a variance of one (1). These two numbers form the random component of the weekly stationary generating process. If this is the first run through the simulation procedure (e.g.,

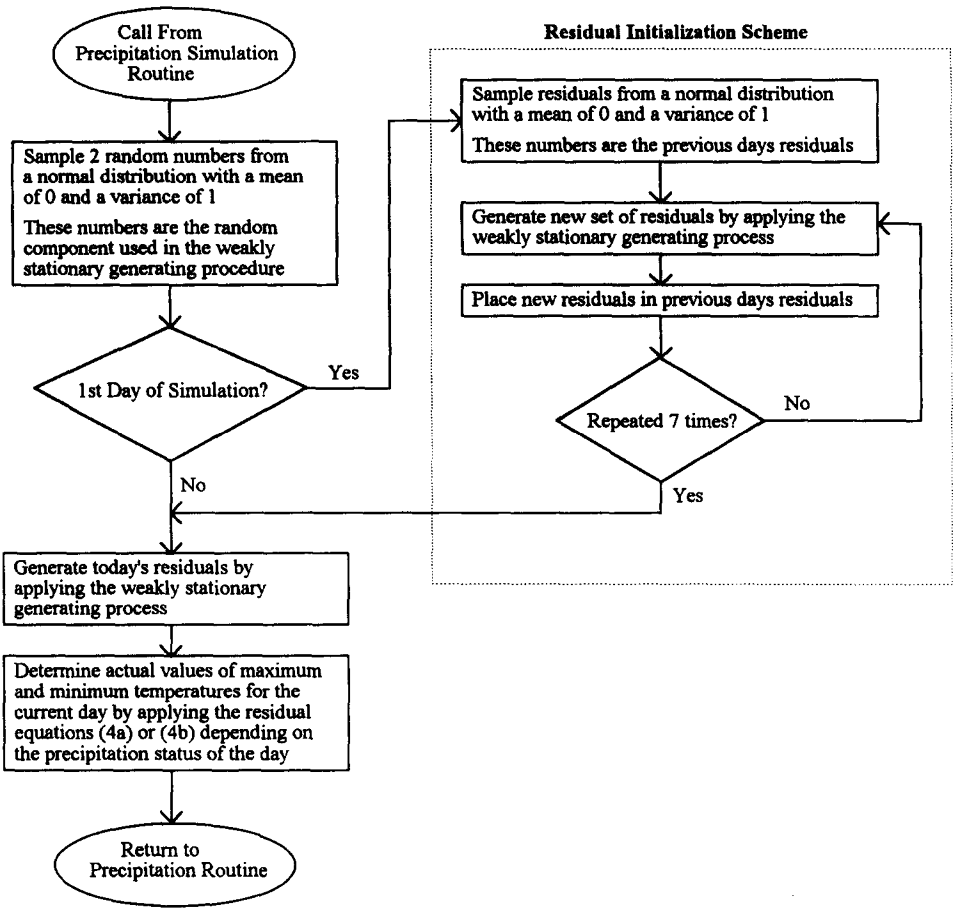

Fig. 9. Temperature simulation procedure.

day 1 of the project) the residual generation procedure is initialized to reduce any transitional startup problems. Therefore, on the first run the previous days residuals are sampled from a normal distribution (mean = 0, variance = 1). A new set of residuals is generated by applying the weekly stationary generating procedure. The new residuals are then placed in variables holding the previous days residuals and the process is repeated. After 7 initializing runs control is passed back to the main routine and the current day's residuals are determined from the weekly stationary generating process. Actual temperature values are determined by solving for $X_{y,d}(i)$ in Eq. (4a) or (4b) depending on the precipitation status of the day. Control is then returned to the precipitation routine.

The model was tested by generating 10 years of weather data and comparing to 20 years of historical weather data from the Edmonton Municipal Airport weather station. In almost all instances, generated weather was agreeable to historical weather data. Tests of underlying assumptions, and testing of results can be found in Appendix A.

6. Sample application

Simulation of a 3-span steel girder bridge construction was undertaken to illustrate the model presented in this paper. The scope of work on this project includes: structural excavation and backfill; placing pipe piling; construction of two abutments and two piers; construction of concrete deck and curbs; rip-rap placement; bridge deck waterproofing; placing asphalt wearing surface; installing bridge rail and guard-rail posts; miscellaneous iron; service ducts; and painting of the pier steel. This project was assumed to start as of June 1, 1994 and to be located in Edmonton region.

A CPM schedule of the project consists of 37 activities. The schedule is based on 10 hour work days, with no provisions for time off for holidays and weekends. Scheduling calculations performed using Microsoft Project [13] revealed that the project will require 111 days of production and will be completed on September 19, 1994.

Prior to simulation, weather sensitive activities should be identified by the modeler. If a group is not specified the activity is considered non-weather sensitive as discussed earlier in the model. Formwork stripping, rip-rap placement, and patching exposed concrete operations are not considered weather sensitive activities as it was felt that they would be delayed only in extreme weather conditions, for example. All other activities were expected to be sensitive to the effects of weather and were included in their respective group where a network was trained for each group individually (details can be found in [22]).

Productivity is predicted based on generated weather conditions using neural networks similar to the one shown in Fig. 10 which is trained on existing productivity data. The networks produce a productivity value when given the weather conditions for the day. Productivity values are essentially the production achieved in a given day under specified conditions divided by the average production for an activity from historical records (e.g. the productivity value used in estimating). Therefore,

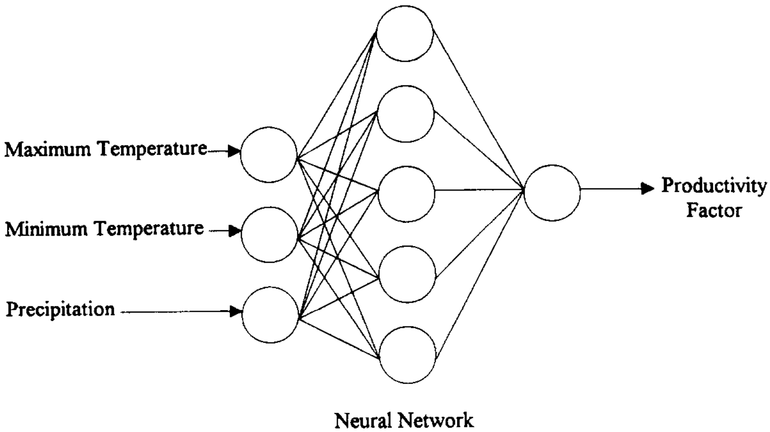

Fig. 10. Productivity neural network.

productivity values can be greater than or less than 1 depending on whether production was better than average or less than average on a particular day.

Since historical productivity data required to support the NN part of the model (for estimating productivity) is not available for this analysis, the NN were trained using productivity data published in the literature that correlate weather to productivity using the IS_NN program. With such assumptions, the results obtained from the simulation will only be as good as the assumptions made and quality of data utilized (same is true of any simulation experiment). This will not be a limitation when the models are applied in a real setting as the contractor would have the required data which are unique for each company (the second author is involved in a study with a major general contractor where the focus is on making productivity data available for estimating purposes in a timely manner). The programs developed in support of network training can be utilized by the contractor to train the NN to the data they have in their historical records. The intent of this sample application is to illustrate how the concepts can be used in a real setting. We have chosen to validate the results of the model in two ways: (1) the weather generation process was tested against data as discussed in Appendix B, and (2) the NN predictions were tested by training the networks using 80% of collected observations and then testing the accuracy by having the networks predict the remaining 20% of the observations and comparing to the actual data. The results were found to be satisfactory and the model to be valid [22].

A simulation model of the project schedule is constructed using a IS_CNV which automatically converts an existing CPM network into the required SLAMSYSTEM simulation code. A section of the resulting simulation model is illustrated in Fig. 11. The model consists of three main components: the project network; detect nodes; and a calendar network. The project network maintains the logic of the project schedule (i.e. CPM structure), detect nodes track the values of continuous variables and determine the time when weather sensitive activities are complete; and the calendar network controls the simulation of weather variables.

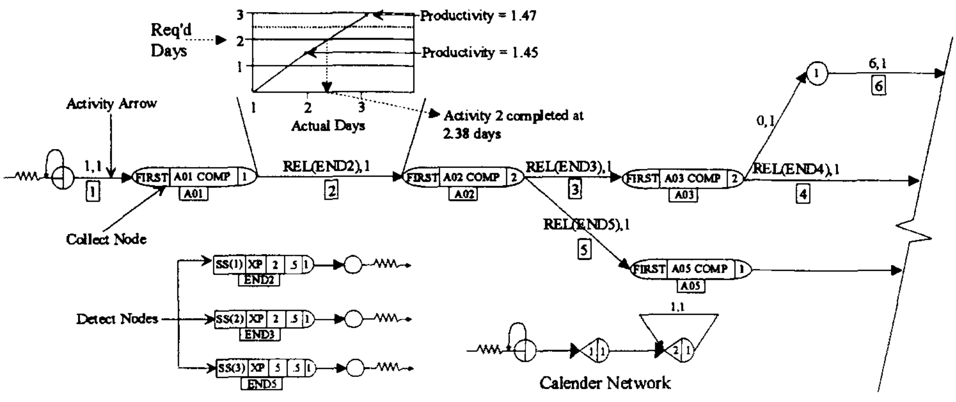

Fig. 11. Project simulation model.

Simulation of the project schedule is performed a number of times in order to have a representative sample of possible project duration and determine confidence intervals for specified project completion times.

It is also often desirable to determine the effect of varying the start date of the project. Project completion times for several start dates are shown in Table 2. Results illustrate the sensitivity of the project schedule to the seasonal weather conditions that will be encountered due to different start dates. April, May, June, or July start dates result in expected project completion times between 103 and 109 days. These estimates show minor variation reflected by the low values obtained for standard deviations and a tight range of possible project completion times. An August start date results in a significant increase in the mean project duration as well as a large increase in the standard deviation indicating a high degree of variability in the project completion time estimate. The range of sample project completion times has increased to 99 days as compared to 15 days with a July 1st start date. The situation gets progressively worse with a September 1st start date as revealed by the increase in mean duration from 138 days to 172 days, a larger standard deviation, and a

Table 2
Effect of change in start date

Start date (1)	Mean duration (days) (2)	Standard deviation (3)	Minimum duration (4)	Maximum duration (5)
April 1, 1994	106	5.27	97	122
May 1, 1994	109	4.71	100	120
June 1, 1994	109	4.14	101	118
July 1, 1994	103	4.09	97	112
August 1, 1994	138	27.0	99	198
September 1, 1994	172	32.6	103	215

range of 112 days. It is evident that the onset of winter weather in a cold climate like Alberta causes significant delay and variability in the project schedule.

7. Conclusions

Project planning can be enhanced using the approach discussed in this paper by planning for weather protection if the analysis reveals that weather will cause delays or planning around the occurrence of bad weather. A further application is in the field of claims analysis. Specifically, claims involving delayed start dates can be analyzed in this manner. It can be shown that delaying the start date by one month in April, May or June has little impact on the project schedule, however, if a delay causes the project to start August 1 rather than July 1 then significant impacts are observed in the project schedule and the contractor may be entitled to compensation depending on the cause of the delay.

The model presented in this paper illustrates a method for generating the occurrence of weather variables, determining the impact on productivity, and modifying activity duration in accordance with the determined impact. The overall process is controlled by the SLAMSYSTEM while weather generation and neural network productivity forecasting is executed in user written FORTRAN code. Therefore, the model can be easily adapted to generate other continuous stochastic processes affecting productivity. The neural networks can also be adapted to determine productivity based on several factors. The model requires the use of SLAMSYSTEM for the simulation phase. In order to assist the modeler and in support of further research in this area we have also automated various tedious tasks including parameterizing the distributions used in the weather generation model, the neural network training and the translation of a PDM network into SLAMSYSTEM. This automation is in the form of executable programs written in Visual Basic requiring a personal computer running Microsoft Windows to work. All developments are in the public domain and are available for further research and developments by contacting the authors.

Appendix A. Test of assumptions

The first implication is that the residual series of each variable is normally distributed. This is tested by calculating the moments of the residual series data and comparing them with the normal distribution. The moments are shown in Table 3. By virtue of the construction the residual series should exhibit a mean of 0 and a standard deviation of 1 [17]. Use of Eqs. (4a) and (4b) successfully reduced the series to residuals with a mean of 0 and a standard deviation of 1. The skewness was slightly negative indicating that the distribution is slightly skewed to the left but it is very close to 0. Kurtosis of the residual series is very close to 3 in both cases which is indicative of a normally distributed sample. For these reasons, it is apparent that the first implication is satisfied.

Table 3
Moments of sample data

Variable (1)	μ (2)	σ (3)	Skewness (4)	Kurtosis (5)
Maximum temperature residuals	0.022	1.089	−0.254	3.002
Minimum temperature residuals	0.037	1.088	−0.314	3.070

The second implication states that the serial correlation of the residual series is described by a first-order linear autoregressive model. A first-order autoregressive model defines serial dependence as $\rho_k = \rho_1^k$, where ρ_k is the serial correlation for lag k. Fig. 12 shows the comparison between the first-order model and the serial correlation coefficients.

It is obvious from Fig. 12 that the serial correlation of the residual series closely approximates that of the first-order linear autoregressive model and therefore the second implication is satisfied. Therefore, this is a viable model for generation of weather variables for project simulation studies.

Appendix B. Output testing

The weather generation model used in this research is tested to ensure that the model produces satisfactory results. Testing is performed by generating ten years of weather data for the Edmonton Municipal Airport Weather station and comparing the simulated means by month to the actual values computed from 20 years of record (1972–1991).

Mean monthly minimum temperatures, maximum temperatures, precipitation amounts, and number of wet days were computed from the actual weather station data. Standard deviations of each of the above parameters was also determined to establish confidence intervals around the calculated means. The simulated means are

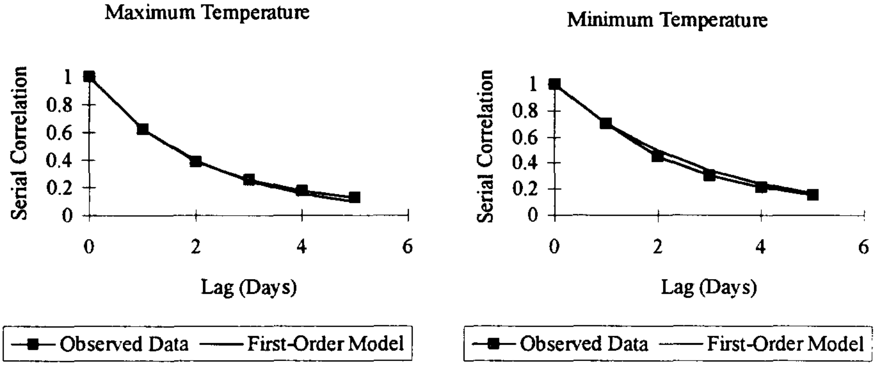

Fig. 12. Autoregressive model vs. serial correlation.

Table 4
Comparison of rainfall amounts

Month (1)	Actual μ (2)	σ (3)	Lower limit (4)	Upper limit (5)	Generated μ (6)	Within limits (7)
January	20.1	15.5	11.30	28.90	22.4	√
February	16.4	8.1	11.80	21.00	17.1	√
March	18.4	12.0	11.59	25.21	17.6	√
April	25.2	12.7	17.99	32.41	33.0	×
May	50.7	29.8	33.78	67.62	54.6	√
June	88.5	42.9	64.14	112.86	84.1	√
July	95.5	51.1	66.49	124.51	100.5	√
August	72.2	39.9	49.55	94.85	84.6	√
September	45.0	28.0	29.10	60.90	47.2	√
October	19.8	16.2	10.60	29.00	20.7	√
November	14.6	11.4	8.13	21.07	13.5	√
December	23.0	13.1	15.56	30.44	21.4	√

then compared to the values obtained from the actual data. For the sake of brevity only the rainfall comparison values are presented. Table 4 lists the actual mean, actual standard deviation, lower and upper limits (1% level) and the mean of the simulated rainfall data. The month of April was the only month to show a significant difference between the actual and generated means for monthly precipitation. All other means were not significantly different at the 1% level. Other generated weather data also exhibited acceptable correlation to the actual weather data.

The model produces satisfactory results for the variables generated in this test. Since very few of the means were significantly different at the 1% level we feel that this weather generating model will operate effectively in the project simulation model.

References

[1] H.N. Ahuja and V. Nandakumar, Simulation model to forecast project completion time, J. Constr. Engrg. Mngmt. **111** (4) (1985) 325–342.

[2] Atmospheric Environment Service, Monthly Record, Environment Canada.

[3] N.B. Benjamin, and T.W. Greenwald, Simulating effects of weather on construction, J. Constr. Engrg. Mngmt. **99** (1973) 175–190.

[4] M.B. Berman, Generating random variates from gamma distributions with noninteger shape parameters, The RAND Corporation, R-641-PR, 1970.

[5] R.I. Carr, Simulation of construction project duration, J. Constr. Div. **105** (1979) 115–125.

[6] G.S. Fishman, Sampling from the gamma distribution on a computer, Comm. ACM **19** (1976) 407–409.

[7] D.W. Halpin, MicroCYCLONE User's Manual (Div. of Constr. Engrg. Mngmt., Purdue University, West Lafayette, IN, 1990).

[8] D.W. Halpin, CYCLONE: Method for modeling of job site processes, J. Constr. Div. **103** (3) (1977) 489–499.

[9] R.B. Harris, Precedence and Arrow Networking Techniques for Construction (John Wiley and Sons, New York, NY, 1978).
[10] R.W. Katz, Probabilistic models, in: A.H. Murphy and R.W. Katz, eds., Probability, Statistics, and Decision Making in the Atmospheric Sciences (Westview Press, Boulder, CO, 1985) 261–285.
[11] M.L. Kavvas, A.A. Aksit and Y.K. Tulanay, A first order nonhomogeneous Markov chain for the daily rainfall occurrences, in: Proc. Symp. on Modeling Hydro. Process (Water Resources Publications, 1977) 44–59.
[12] N.C. Matalas, Mathematical assessment of synthetic hydrology, Water Resources Res. **3** (4) (1967) 937–45.
[13] Microsoft Project for Windows (Microsoft Corporation, 1992).
[14] NeuroWindows – Neural Network Dynamic Link Library (Ward Systems Group, Frederick, MD, 1991).
[15] A.A.B. Pritsker, Introduction to Simulation and SLAMSYSTEM (Halsted Press, New York, NY, 3rd ed., 1986).
[16] A. Pritsker, C.E. Sigal and R.D.J. Hammesfahr, SLAMSYSTEM Network Models for Decision Support (Prentice Hall, Englewood Cliffs, NJ, 1989).
[17] C.W. Richardson, Stochastic simulation of daily precipitation, temperature, and solar radiation, Water Resources Res. **17** (1) (1981) 182–90.
[18] SLAMSYSTEM (Pritsker Corporation, Indianapolis, IN, 1990).
[19] G.R. Smith and D.E. Hancher, Estimating precipitation impacts for scheduling, J. Constr. Engrg. Mngmt. **115** (4) (1989) 552–566.
[20] P.R. Tadikamalla, Computer generation of gamma random variables, Comm. ACM **21** (1978) 419–21.
[21] P.R. Tadikamalla, Computer generation of gamma random variables II, Comm. ACM **21** (1978) 925–27.
[22] R.J. Wales, An integrated model for construction project simulation, Unpublished M.Sc. Thesis, Department of Civ. Engineering, University of Alberta.
[23] V.Y. Yevjevich, Structural analysis of hydrologic time series, Hydrol. Pap. **56** (1972) 59 pp.
[24] G.K. Young, Discussion of mathematical assessment of synthetic hydrology, in: N.C. Matalas, Water Resources Res. **4** (3) (1968) 681–682.