Testing of a Monthly Weather Model[1]

Rafael F. Dávila[2]

ABSTRACT

A mathematical model which can be used to generate monthly time series of precipitation and temperature that resemble the historic weather records in terms of their mean, variance, skewness and month to month correlations was tested. The basic model is the lag one Markov model with a Gamma distributed random variable.

The model was tested in three locations within the continental United States (Omaha, Nebraska; Ithaca, New York; and San Francisco, California) and in one location in Puerto Rico (San Juan). Test results indicate that the model does very well in reproducing the first two moments, mean and variance, of the historic records when both precipitation and temperature are generated. The model was capable of reproducing the coefficient of skewness for 50% of the records at the four locations tested. The lag one monthly coefficient of correlation was well reproduced for 65% of the records.

Although the work on the model was motivated by the need for synthetic precipitation and temperature data for the preliminary design of land treatment systems for the disposal of wastewater, its use is not limited to this type of project.

INTRODUCTION

Long sequences of precipitation and temperature are often needed for environmental planning. The level of detail of the planning effort determines the time step of the sequences. For very detailed planning, records with a time step as small as an hour might be needed; in other projects, a yearly time step might suffice. For preliminary design of land treatment systems a monthly time step is appropriate (8).

When long periods of historical records are available, these records can be divided into sequences of length equal to the planning horizon and used to simulate the system's response. In land treatment of wastewater projects the planning horizon is at least 25 years; therefore, for most locations the historical records allow only for one or two sequences. To overcome this difficulty, the planner or engineer may employ simulated or synthetic sequences.

Synthetic sequences or time series of monthly precipitation and temperature that have the statistical characteristics of the available historical records can be constructed. For monthly time series, the most meaningful characteristics of the historical data that need to be reproduced are the mean, variance, skewness and lag-one autocorrelation between monthly values. A basic assumption used in generating these sequences is that

[1] Manuscript submitted to Editorial Board May 17, 1985.

[2] Agricultural Engineer, Agricultural Extension Service, Mayagüez Campus, University of Puerto Rico.

the processes are stationary; that is, that the statistical characteristics of the records will remain the same over time (2).

A mathematical model that can be used to generate monthly time series of precipitation and temperature was tested. This model was used to generate synthetic data for three locations within the continental United States and for one location in Puerto Rico. The statistical parameters of the generated time series were compared with the statistical parameters of the historical data in order to justify the model's use. Although the work on the model was motivated by the need for synthetic precipitation and temperature data for the preliminary design of land treatment systems for the disposal of wastewater, its use is not limited to this type of project (1).

MATHEMATICAL ASSESSMENT

The basic model used to generate the monthly sequences is the lag-one Markov model, which is represented as

$$(x_t - \mu_t)/\sigma_t = \rho_t(x_{t-1} - \mu_{t-1})/\sigma_{t-1} + \epsilon(1 - \rho_t^2)^{1/2} \qquad /1/$$

where, x represents the precipitation amount or temperature, μ_t is the mean, σ_t is the standard deviation, ρ_t the correlation with the previous month, and ϵ is a random component with zero mean and unit variance. The subscript t refers to the month.

In order to generate synthetic time series that resemble the historic events in terms of the mean, variance, coefficient of skewness and correlation with the previous month, μ_t, σ_t and ρ_t must be by their estimators, $\hat{\mu}_t$, $\hat{\sigma}_t$ and $\hat{\rho}_t$ (3). In addition, the random component must consider skewness. The random component ϵ may be computed as

$$\epsilon = (2/\gamma_\epsilon)[1 + (\gamma_\epsilon \cdot K/6) - (\gamma_\epsilon^2/36)]^3 - (2/\gamma_\epsilon) \qquad /2/$$

where, the skewness of ϵ, γ_ϵ is related to the coefficient of skewness in month t and month $t-1$, $\hat{\gamma}_t$ and $\hat{\gamma}_{t-1}$ respectively (3, 4). This relationship is described by

$$\gamma_\epsilon = (\hat{\gamma}_t - \hat{\rho}_t^3\hat{\gamma}_{t-1})/(1 - \hat{\rho}_t^3)^{3/2} \qquad /3/$$

The variable K in equation /2/ is a standard normal variate. The random component, ϵ defined this way is a gamma distributed variate with zero mean and unit variance and coefficient of skewness γ_ϵ (3).

Incorporating the modifications described above and rearranging the terms, equation /1/ can be expressed as

$$x_t = [\hat{\rho}_t(x_{t-1} - \hat{\mu}_{t-1})/\hat{\sigma}_{t-1} + \epsilon(1 - \hat{\rho}_t^2)^{1/2}]\ \hat{\sigma}_t + \hat{\mu}_t \qquad /4/$$

This model was first proposed by Thomas and Fiering (7) for the generation of streamflow sequences, but as will be justified in the next section, the model can be used to generate monthly sequences of precipitation and temperature with acceptable results. The model has been used successfully to generate monthly precipitation sequences in the Quae Yai river basin in Thailand (4).

The procedure used to generate a time series of monthly precipitation or temperature of length L using equations /2/ to /4/ is detailed in the following outline:

1. First, analyze the available historical records for monthly mean, variance, coefficient of skewness and lag-one correlation between months.
2. Calculate the coefficient of skewness of the random component, γ_ϵ for each month using equation /3/.
3. Generate a set of standard normal variates of length L and use equation /2/ to calculate an equivalent set of gamma random variates.
4. Using an initial guess for x_{t-1}, using equation /4/ to produce the synthetic time series. After the first 12 values have been generated, $\hat{\mu}_{t-1}$, $\hat{\sigma}_{t-1}$ and $\hat{\gamma}_{t-1}$ take the initial values again. Negative precipitation values should be set to zero.
5. If more than one time series is desired, steps 2 through 4 are repeated.
6. In order to minimize the bias associated with the initial guess, at least the first five time series generated using this procedure should be discarded.

MODEL TESTING

In order to test the weather model given by equation /4/ precipitation and temperature data were collected for four locations: Ithaca, NY; San Juan, PR; Omaha, NE; and San Francisco, CA. These data were analyzed, synthetic precipitation and temperature time series generated, and the results evaluated on the basis of how well the statistical parameters of the historic data were reproduced. The evaluation of the results was done with statistical tests which are described below in statistical methods.

The use of a lag-one Markov model was justified by testing the significance of the lag-one monthly correlations of precipitation and temperature. The use of a gamma random variable was also justified; this was done by testing the significance of the coefficient of skewness. A normal distribution of precipitation or temperature has a skewness coefficient of zero; therefore if all the monthly skewness coefficients turn out nonsignificant, the gamma random variable could be substituted for a normal random variable, meaning that the weather parameters are normally distributed. If at least one of the monthly coefficients of skewness is significant, a gamma random variable is more appropriate.

Statistical methods

The significance of the coefficient of correlation was tested with a *t*-statistic as described by Snedecor and Cochran (5). The following hypotheses were established:

$$H_0: \{\hat{\rho}_t = 0\}$$

$$H_A: \{\hat{\rho}_t \neq 0\}$$

The term $\hat{\rho}_t$ = the lag-one coefficient of correlation.

To test the coefficient of skewness, the following hypotheses were checked:

$$H_0: \{\hat{\gamma}_t = 0\}$$

$$H_A: \{\hat{\gamma}_t \neq 0\}$$

where, $\hat{\gamma}_t$ = the coefficient of skewness. A *t* statistic was also used to test $\hat{\gamma}_t$. The test is described by Snedecor and Cochran (5).

The generated data were evaluated by testing the significance of the difference between the historic and generated mean, variances, coefficient of skewness and coefficient of correlation. Average monthly values of 20 time series of 20 years were used to test against the monthly averages of the historic records for these parameters.

The following general hypotheses were tested:

$$H_0: \{X_H = X_G\}$$

$$H_A: \{X_H \neq X_G\}$$

The X term refers to the parameter being tested, and the subscripts H and G are the historic and generated valued of X, respectively. Statistical tests from Steel and Torrie (6) were used for the means and variances. The coefficients of correlation were tested with the test proposed by Snedecor and Cochran (5).

In order to test the coefficients of skewness, I modified the test proposed by Snedecor and Cochran (5) for a single sample so that the equality of two samples could be tested. The difference, D, between the two samples is assumed to be normally distributed. The difference between the historic and generated coefficients of skewness is given by equation /5/ and their variance, σ_D, by /6/.

$$D = Y_H - Y_G \qquad /5/$$

$$\sigma_D = \sqrt{Y^2_{H/n_H} + Y^2_{G/n_G}} \qquad /6/$$

Analysis of historic weather records

The statistical parameters of the historic data are given in tables 1 through 8. The results of the significance test on the coefficients of correlation are shown in table 9. Table 9 also shows that for the four

TABLE 1.—*Statistical analysis of precipitation for Ithaca, New York*[1]

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[2]
	cm	*cm*^{2}	*cm*		
J	5.17	9.70	3.11	1.96	0.19
F	5.37	7.41	2.72	0.11	−0.02
M	6.36	7.13	2.67	0.60	−0.04
A	7.12	5.70	2.59	1.05	0.04
M	7.64	8.28	2.88	0.12	−0.01
J	9.46	23.32	4.83	2.35	0.31
J	9.02	19.02	4.36	0.52	−0.03
A	8.61	10.07	3.17	0.25	0.05
S	8.72	19.80	4.45	1.40	0.16
O	8.29	26.10	5.11	0.94	0.21
N	7.51	12.76	3.57	0.66	−0.12
D	6.60	8.00	2.83	0.05	0.26

[1] Based on 31 years of records.
[2] Correlation with previous month.

TABLE 2.—*Statistical analysis of precipitation for San Juan, Puerto Rico*[1]

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[2]
	cm	*cm*^{2}	*cm*		
J	8.08	17.75	4.21	1.11	0.10
F	5.58	16.83	4.10	1.78	0.17
M	5.83	12.85	3.59	0.79	0.03
A	9.08	33.14	5.76	0.58	0.34
M	14.85	82.61	9.09	0.71	0.03
J	12.50	43.73	6.61	0.57	0.59
J	12.64	35.46	5.96	0.63	0.45
A	15.45	26.46	5.14	0.65	0.25
S	15.21	37.91	6.16	0.60	0.05
O	15.11	43.61	6.60	1.24	−0.08
N	14.49	55.13	7.43	0.77	0.32
D	12.15	37.43	6.12	0.78	0.23

[1] Based on 28 years of records.
[2] Correlation with previous month.

locations tested, the month to month dependence of precipitation and temperature cannot be ignored in the general weather model. The results also indicate that the month to month correlation is more important in San Juan and San Francisco, which are the two locations with warmer

TABLE 3.—*Statistical analysis of precipitation for Omaha, Nebraska**

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[2]
	cm	*cm*^{2}	*cm*		
J	1.94	2.70	1.64	2.55	−0.02
F	2.55	3.37	1.84	0.83	−0.03
M	4.28	8.47	2.91	1.49	0.01
A	7.50	15.82	3.98	0.41	0.12
M	10.44	28.06	5.30	0.46	0.16
J	11.19	29.79	5.46	0.78	0.09
J	9.16	29.29	5.41	0.71	−0.34
A	9.57	31.80	5.64	0.61	−0.14
S	8.26	39.72	6.30	1.90	−0.14
O	5.14	13.90	3.73	0.45	−0.09
N	3.29	7.63	2.76	0.69	0.23
D	2.14	2.90	1.70	1.33	0.18

[1] Based on 40 years of records.
[2] Correlation with previous month.

TABLE 4.—*Statistical analysis of precipitation for San Francisco, California*[1]

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[2]
	cm	*cm*^{2}	*cm*		
J	11.13	48.98	7.00	0.75	0.30
F	7.55	27.95	5.29	0.90	0.30
M	7.29	20.32	4.51	0.61	0.29
A	3.82	13.50	3.67	1.08	0.40
M	1.22	2.05	1.43	2.76	0.09
J	0.35	0.48	0.69	3.06	−0.15
J	0.09	0.06	0.25	5.23	−0.06
A	0.17	0.15	0.39	3.42	−0.09
S	0.60	1.19	1.09	2.71	0
O	2.79	11.45	3.38	1.85	0.06
N	6.49	25.42	5.04	0.78	0.15
D	9.64	43.76	6.62	1.05	0.27

[1] Based on 40 years of records.
[2] Correlation with previous month.

TABLE 5.—*Statistical analysis of temperature for Ithaca, New York*[1]

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[2]
	°C	°C²	°C		
J	−5.60	5.90	2.43	−0.10	0.38
F	−5.03	7.26	2.70	−0.24	−0.05
M	0.12	4.22	2.05	−0.41	−0.05
A	6.96	2.98	1.73	−0.61	0.26
M	12.67	2.61	1.62	−0.24	−0.10
J	17.93	1.22	1.11	−0.04	−0.32
J	20.41	1.12	1.06	1.01	0.07
A	19.49	1.14	1.07	0.89	0.48
S	15.63	2.00	1.42	0.40	0.27
O	9.68	3.13	1.77	0.04	0.18
N	4.01	2.77	1.66	−0.40	0.19
D	−2.53	4.86	2.21	−0.46	0.15

[1] Based on 31 years of records.
[2] Correlation with previous month.

TABLE 6.—*Statistical analysis of temperature for San Juan, Puerto Rico*[1]

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[2]
	°C	°C²	°C		
J	24.53	0.43	0.65	0.37	0.66
F	24.55	0.55	0.74	0.12	0.90
M	25.09	0.57	0.76	−0.27	0.86
A	25.81	0.42	0.65	−0.04	0.81
M	26.68	0.43	0.66	0.08	0.78
J	27.39	0.45	0.67	0.22	0.83
J	27.59	0.39	0.62	0.46	0.77
A	27.77	0.21	0.46	0.42	0.88
S	27.61	0.31	0.55	0.65	0.85
O	27.28	0.36	0.60	0.14	0.79
N	26.32	0.37	0.61	0.14	0.86
D	25.21	0.43	0.65	−0.01	0.72

[1] Based on 28 years of records.
[2] Correlation with previous month.

TABLE 7.—*Statistical analysis of temperature for Omaha, Nebraska*[1]

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[2]
	°C	°C²	°C		
J	−5.86	7.97	2.82	−0.12	−0.11
F	−2.65	8.22	2.87	0.23	0.01
M	2.97	8.70	2.95	−0.07	0.35
A	11.19	3.76	1.94	0.18	0.14
M	17.26	3.88	1.97	0.45	0.07
J	22.58	2.47	1.57	−0.01	0.08
J	25.30	1.97	1.40	0.16	0.10
A	24.02	2.17	1.47	0.37	0.04
S	18.88	2.66	1.63	−0.09	0.30
O	13.02	3.94	1.99	0.35	0.10
N	4.09	3.18	1.78	−0.32	0.28
D	−2.53	6.46	2.54	0.07	0.16

[1] Based on 40 years of records.
[2] Correlation with previous month.

TABLE 8.—*Statistical analysis of temperature for San Francisco, California*[1]

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[2]
	°C	°C²	°C		
J	10.56	1.55	1.25	−0.46	0.46
F	11.94	1.33	1.15	0.02	0.28
M	12.39	1.12	1.06	0.86	0.39
A	12.94	1.03	1.02	−0.44	0.22
M	13.72	0.75	0.87	0.72	0.05
J	14.78	0.78	0.89	0.64	0.27
J	14.82	0.42	0.65	−0.29	0.44
A	15.35	0.67	0.82	−0.11	0.31
S	16.85	1.15	1.07	0.17	0.35
O	16.34	0.84	0.92	0.29	0.16
N	14.02	0.90	0.95	0.54	0.28
D	11.15	1.75	1.32	0	0.43

[1] Based on 40 years of records.
[2] Correlation with previous month.

TABLE 9.—*Test for significance (10% level) on coefficient of correlation*

Months	Ithaca	San Juan	Omaha	San Francisco
		Precipitation		
D-J	NS[1]	NS	NS	NS
J-F	NS	NS	NS	S[2]
F-M	NS	NS	NS	S
M-A	NS	S	NS	S
A-M	NS	NS	NS	NS
M-J	S	S	NS	NS
J-J	NS	S	S	NS
J-A	NS	NS	NS	NS
A-S	NS	NS	NS	NS
S-O	NS	NS	NS	NS
O-N	NS	S	S	NS
N-D	NS	NS	NS	S
		Temperature		
D-J	NS	S	NS	NS
J-F	NS	S	NS	S
F-M	NS	S	S	S
M-A	NS	S	NS	NS
A-M	NS	S	NS	NS
M-J	S	S	NS	S
J-J	NS	S	NS	S
J-A	S	S	NS	S
A-S	NS	S	S	S
S-O	NS	S	NS	NS
O-N	NS	S	S	S
N-D	NS	S	NS	S

[1] Nonsignificant.
[2] Significant.

weather. Even though only four locations were tested, it is believed that the use of a lag-one Markov model to generate synthetic precipitation and temperature on a monthly basis is justified.

The test on the coefficient of skewness is given in table 10. These results indicate that monthly precipitation and temperature cannot be assumed to be normally distributed in the general weather model. The distribution of monthly precipitation had significant skewness in all

TABLE 10.—*Test for significance (10% level) on coefficient of skewness*

Month	Ithaca	San Juan	Omaha	San Francisco
		Precipitation		
J	S[1]	S	S	S
F	S	S	S	S
M	S	S	S	S
A	S	S	S	S
M	S	S	S	S
J	S	S	S	S
J	S	S	S	S
A	S	S	S	S
S	S	S	S	S
O	S	S	S	S
N	S	S	S	S
D	S	S	S	S
		Temperature		
J	S	S	S	S
F	S	S	S	NS[1]
M	S	S	S	S
A	S	NS	S	S
M	S	S	S	S
J	NS	S	NS	S
J	S	S	S	S
A	S	S	S	S
S	S	S	S	S
O	NS	S	S	S
N	S	S	S	S
D	S	NS	S	NS

[1] See table 9.

months and locations tested. The distribution of monthly temperature also had significant skewness in most of the records tested. On the basis of these results, a gamma random variable seems appropriate for the synthetic generation of monthly precipitation and temperature.

GENERATION OF SYNTHETIC WEATHER RECORDS

With the procedure outlined in Mathematical Assessment twenty time series of length equal to 20 years (240 months) were generated for Ithaca, New York; San Juan, Puerto Rico; Omaha, Nebraska; and San Francisco, California. The statistics for each time series were computed independently, and then the average for the 20 series was computed and reported. In order to smooth the bias associated with initial values, five time series were initially run through at each location and not used to compute the average statistics.

TABLE 11.—*Statistical analysis of synthetic precipitation for Ithaca, New York*

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[1]
	cm	*cm*^{2}	*cm*		
J	5.36	11.36	3.37	1.44	0.21
F	5.40	6.90	2.63	0.21	−0.01
M	6.51	6.75	2.60	0.27	0.08
A	7.01	5.95	2.44	0.93	0.11
M	7.83	6.36	2.52	0.15	−0.05
J	9.19	18.39	4.29	1.37	0.34
J	8.82	19.62	4.43	0.41	−0.06
A	8.53	10.77	3.28	0.29	0.15
S	8.84	19.22	4.38	0.81	0.23
O	8.56	23.26	4.82	0.63	0.20
N	7.89	12.59	3.55	0.35	−0.11
D	6.70	7.99	2.83	−0.10	0.28

[1] Correlation with previous month.

TABLE 12.—*Statistical analysis of synthetic precipitation for San Juan, Puerto Rico*

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[1]
	cm	*cm*^{2}	*cm*		
J	8.08	15.29	3.91	0.63	0.11
F	5.54	15.01	3.87	1.01	0.23
M	5.64	12.32	3.51	0.55	0.09
A	9.08	31.04	5.57	0.40	0.35
M	15.67	88.45	9.40	0.56	0.02
J	13.13	39.89	6.32	0.25	0.62
J	12.75	32.06	5.66	0.44	0.43
A	15.26	23.40	4.84	0.22	0.32
S	15.01	38.45	6.20	0.62	0.03
O	15.33	39.73	6.30	0.94	0.02
N	14.43	48.25	6.95	0.48	0.30
D	11.82	30.73	5.54	0.37	0.25

[1] See table 11.

RESULTS AND DISCUSSION

Although many values of historic correlation coefficient and skewness coefficient are so low that they are statistically insignificant, they were not set to zero in the input data set for the synthetic series. The results for the synthetic series are given in tables 11 to 18.

Comparison of synthetic and historic weather records

The historic and generated weather records were compared by means of the statistical tests outlined in Statistical Methods. The differences between the average historical and generated parameters were tested for significance at the 10% level.

TABLE 13.—*Statistical analysis of synthetic precipitation for Omaha, Nebraska*

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[1]
	cm	*cm*^{2}	*cm*		
J	1.84	2.39	1.55	1.60	0.06
F	2.62	2.89	1.70	0.56	−0.08
M	4.24	8.33	2.89	1.20	0.01
A	7.53	14.41	3.80	0.20	0.08
M	10.36	24.29	4.93	0.35	0.05
J	11.20	28.28	5.32	0.57	0.02
J	9.62	27.71	5.26	0.57	−0.29
A	9.17	27.94	5.29	0.55	−0.12
S	7.89	32.65	5.71	1.23	−0.11
O	4.98	12.10	3.48	0.42	0.09
N	3.54	6.70	2.59	0.61	0.21
D	2.10	2.56	1.60	0.93	0.28

[1] See table 11.

TABLE 14.—*Statistical analysis of synthetic precipitation for San Francisco, California*

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[1]
	cm	*cm*^{2}	*cm*		
J	10.89	42.63	6.54	0.52	0.29
F	7.40	27.68	5.26	0.97	0.37
M	7.47	18.65	4.32	0.50	0.31
A	4.24	12.51	3.54	0.88	0.33
M	1.16	1.68	1.30	1.65	0.07
J	0.35	0.29	0.54	1.90	−0.03
J	0.07	0.04	0.20	2.76	−0.05
A	0.18	0.14	0.37	2.26	−0.02
S	0.63	1.04	1.02	1.69	−0.06
O	2.72	9.89	3.14	1.40	0.04
N	6.87	23.07	4.80	0.53	0.21
D	9.87	35.06	5.92	0.67	0.24

[1] See table 11.

The means of the historical precipitation and temperature records were very well reproduced. No significant difference was found between historic and generated means for any month or location.

The variances of the historical precipitation and temperature records were also well reproduced. No significant difference was found between historic and generated values except for the months of June and July in San Francisco. Both of these variances had very low values and were associated with extremely low average precipitation, 0.35 cm and 0.07 cm for June and July respectively.

The historic and generated coefficients of skewness are presented in

TABLE 15.—*Statistical analysis of synthetic temperature for Ithaca, New York*

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[1]
	°C	°C²	°C		
J	−5.50	6.18	2.49	0.01	0.40
F	−5.03	7.13	2.67	−0.20	−0.04
M	0.23	4.14	2.03	−0.51	−0.07
A	6.92	2.69	1.64	−0.25	0.29
M	12.80	2.03	1.42	−0.06	−0.18
J	17.82	1.17	1.08	−0.19	−0.30
J	20.36	1.14	1.07	0.76	0.02
A	19.46	1.32	1.15	0.72	0.54
S	15.65	2.02	1.42	0.24	0.32
O	9.78	2.80	1.67	0.02	0.17
N	4.21	2.59	1.61	−0.36	0.19
D	−2.49	5.04	2.24	−0.56	0.15

[1] See table 11.

TABLE 16.—*Statistical analysis of synthetic temperature for San Juan, Puerto Rico*

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[1]
	°C	°C²	°C		
J	24.58	0.43	0.66	0.10	0.56
F	24.60	0.60	0.77	−0.04	0.91
M	25.15	0.58	0.76	−0.16	0.85
A	25.83	0.39	0.63	0.03	0.80
M	26.73	0.37	0.61	0	0.79
J	27.39	0.38	0.62	0.16	0.79
J	27.57	0.33	0.58	0.52	0.69
A	27.76	0.20	0.45	0.32	0.85
S	27.61	0.32	0.56	0.44	0.85
O	27.30	0.34	0.59	0.18	0.77
N	26.37	0.34	0.58	0.15	0.84
D	25.25	0.40	0.63	0	0.66

[1] See table 11.

table 19. The results of testing the significance of the differences between these parameters are given in table 20. These results are not as good as those for the means and variances. Fifty-five percent of the precipitation records showed no significant differences between historic and generated values. This percentage was 50% for the temperature records. The

TABLE 17.—*Statistical analysis of synthetic temperature for Omaha, Nebraska*

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[1]
	°C	°C²	°C		
J	−6.02	8.14	2.85	−0.11	−0.04
F	−2.75	7.49	2.74	0.04	−0.03
M	2.90	8.26	2.87	0.05	0.36
A	11.25	3.82	1.95	0.26	0.07
M	17.32	3.71	1.93	0.42	0.05
J	22.68	2.74	1.65	0.07	0.02
J	25.32	1.90	1.38	−0.07	0.07
A	23.97	1.95	1.40	1.32	0.06
S	18.87	2.45	1.57	−0.01	0.32
O	13.03	3.69	1.92	0.21	0.13
N	4.08	3.23	1.80	−0.26	0.29
D	−2.50	5.86	2.42	−0.08	0.17

[1] See table 11.

TABLE 18.—*Statistical analysis of synthetic temperature for San Francisco, California*

Month	Mean	Variance	Std. Dev.	Skew Coeff.	Corr. Coeff.[1]
	°C	°C²	°C		
J	10.58	1.53	1.24	−0.52	0.39
F	11.86	1.37	1.17	0.06	0.22
M	12.36	1.07	1.04	0.45	0.35
A	12.92	1.06	1.03	−0.23	0.25
M	13.69	0.67	0.82	0.60	0.11
J	14.78	0.74	0.86	0.45	0.23
J	14.83	0.43	0.65	−0.13	0.49
A	15.39	0.69	0.83	0.03	0.33
S	16.81	0.94	0.97	0.07	0.30
O	16.30	0.76	0.87	0.22	0.07
N	14.08	0.83	0.91	0.15	0.30
D	11.15	1.55	1.25	−0.18	0.34

[1] See table 11.

TABLE 19.—*Historic and generated coefficients of skewness*

Month	Ithaca		San Juan		Omaha		San Francisco	
	Historic	Generated	Historic	Generated	Historic	Generated	Historic	Generated
				Precipitation				
J	1.96	1.44	1.11	0.63	2.55	1.60	0.75	0.52
F	0.11	0.21	1.78	1.01	0.83	0.56	0.90	0.97
M	0.06	0.27	0.79	0.55	1.49	1.20	0.61	0.50
A	1.05	0.93	0.58	0.40	0.41	0.20	1.08	0.88
M	0.12	0.15	0.71	0.56	0.46	0.35	2.76	1.65
J	2.35	1.37	0.57	0.25	0.78	0.57	3.06	1.90
J	0.52	0.41	0.63	0.44	0.71	0.57	5.23	2.76
A	0.25	0.29	0.65	0.22	0.61	0.55	3.42	2.26
S	1.40	0.81	0.60	0.62	1.90	1.23	2.71	1.69
O	0.94	0.63	1.24	0.94	0.45	0.42	1.85	1.40
N	0.66	0.35	0.77	0.48	0.69	0.61	0.78	0.53
D	0.05	−0.10	0.78	0.37	1.33	0.93	1.05	0.67
				Temperature				
J	−0.10	0.01	0.37	0.10	−0.12	−0.11	−0.46	−0.52
F	−0.24	−0.20	0.12	−0.04	0.23	0.04	0.02	0.06
M	−0.41	−0.51	−0.27	−0.16	−0.07	0.05	0.86	0.45
A	−0.61	−0.25	−0.04	0.03	0.18	0.26	−0.44	−0.23
M	−0.24	−0.06	0.08	0.0	0.45	0.42	0.72	0.60
J	−0.04	−0.19	0.22	0.16	−0.01	0.07	0.64	0.45
J	1.01	0.76	0.46	0.52	0.16	−0.07	−0.29	−0.13
A	0.89	0.72	0.42	0.32	0.37	1.32	−0.11	0.03
S	0.40	0.24	0.65	0.44	−0.09	−0.01	0.17	0.07
O	0.04	0.02	0.14	0.18	0.35	0.21	0.29	0.22
N	−0.40	−0.36	0.14	0.15	−0.32	−0.26	0.54	0.15
D	−0.46	−0.56	−0.01	0.0	0.07	−0.08	0.0	−0.18

TABLE 20.—*Test for significance (10% level) of the differences between historical and generated skewness*

Month	Ithaca	San Juan	Omaha	San Francisco
		Precipitation		
J	NS[1]	S[2]	S	NS
F	S	S	S	NS
M	S	NS	NS	NS
A	NS	NS	S	NS
M	NS	NS	NS	S
J	S	S	NS	S
J	NS	NS	NS	S
A	NS	S	NS	S
S	S	NS	S	S
O	NS	NS	NS	NS
N	S	S	NS	NS
D	S	S	NS	S
		Temperature		
J	S	S	NS	NS
F	NS	S	S	S
M	NS	S	S	S
A	S	S	S	S
M	S	S	NS	NS
J	S	NS	S	NS
J	NS	NS	S	NS
A	NS	NS	NS	S
S	S	NS	NS	S
O	NS	NS	S	NS
N	NS	NS	NS	S
D	NS	S	NS	S

[1]Nonsignificant.
[2] Significant.

TABLE 21.—*Historic and generated coefficients of correlation*

Month	Ithaca		San Juan		Omaha		San Francisco	
	Historic	Generated	Historic	Generated	Historic	Generated	Historic	Generated
				Precipitation				
J	0.19	0.21	0.10	0.11	−0.02	0.06	0.30	0.29
F	−0.02	−0.01	0.17	0.23	−0.03	−0.08	0.30	0.37
M	−0.04	0.08	0.03	0.09	0.01	0.01	0.29	0.31
A	0.04	0.11	0.34	0.35	0.12	0.08	0.40	0.33
M	−0.01	−0.05	0.03	0.02	0.16	0.05	0.09	0.07
J	0.31	0.34	0.59	0.62	0.09	0.02	−0.15	−0.03
J	−0.03	−0.06	0.45	0.43	−0.34	−0.29	−0.06	−0.05
A	0.05	0.15	0.25	0.32	−0.14	−0.12	−0.09	−0.02
S	0.16	0.23	0.05	0.03	−0.14	−0.11	0	−0.06
O	0.21	0.20	−0.08	0.02	−0.09	0.09	0.06	0.04
N	−0.12	−0.11	0.32	0.30	0.23	0.21	0.14	0.21
D	0.26	0.20	0.23	0.25	0.18	0.28	0.27	0.24
				Temperature				
J	0.38	0.40	0.66	0.56	−0.11	−0.04	0.46	0.39
F	−0.05	−0.04	0.90	0.91	0.01	−0.03	0.28	0.22
M	−0.05	−0.07	0.86	0.85	0.35	0.36	0.39	0.35
A	0.26	0.29	0.81	0.80	0.14	0.07	0.22	0.25
M	−0.10	−0.18	0.78	0.79	0.07	0.05	0.05	0.11
J	−0.32	−0.30	0.83	0.79	0.08	0.02	0.27	0.23
J	0.07	0.02	0.77	0.69	0.10	0.07	0.44	0.49
A	0.48	0.54	0.88	0.85	0.04	0.06	0.31	0.33
S	0.27	0.32	0.85	0.85	0.30	0.32	0.35	0.30
O	0.18	0.17	0.79	0.77	0.10	0.13	0.16	0.07
N	0.19	0.19	0.86	0.84	0.28	0.29	0.28	0.30
D	0.15	0.15	0.72	0.66	0.16	0.17	0.43	0.34

differences between historic and generated records for precipitation might in part be due to the truncation of negative values.

The historic and generated coefficients of correlation are given in table 21. Similar test results for these parameters are shown in table 22. The coefficient of correlation was well reproduced for 65% of the precipitation records and 71% of the temperature records.

TABLE 22.—*Test for significance (10% level) of the differences between historical and generated coefficients of correlation*

Months	Ithaca	San Juan	Omaha	San Francisco
		Precipitation		
D-J	NS[1]	NS	S[2]	NS
J-F	NS	NS	S	S
F-M	NS	NS	NS	NS
M-A	NS	NS	NS	S
A-M	NS	NS	S	NS
M-J	NS	NS	S	S
J-J	NS	NS	NS	NS
J-A	S	NS	NS	S
A-S	S	NS	NS	S
S-O	NS	S	NS	NS
O-N	NS	NS	NS	S
N-D	NS	NS	S	NS
		Temperature		
D-J	NS	S	S	S
J-F	NS	NS	NS	S
F-M	NS	NS	NS	NS
M-A	NS	NS	S	NS
A-M	S	NS	NS	NS
M-J	NS	S	S	NS
J-J	NS	S	NS	S
J-A	S	S	NS	NS
A-S	NS	NS	NS	S
S-O	NS	NS	NS	S
O-N	NS	NS	NS	NS
N-D	NS	S	NS	S

[1] Nonsignificant.
[2] Significant.

Conclusions

The use of the lag one model was justified by testing the significance of the lag one monthly correlation coefficients of precipitation and temperature. The use of gamma distributed random variable was also justified against using a more simple normally distributed random variable. This was done by testing the significance of the coefficient of skewness of the distributions of precipitation and temperature. The model was tested at three locations in the continental United States and one location in Puerto Rico.

The results of testing the model indicate that the model does very well in reproducing the first two moments, mean and variance, of the historic records when both precipitation and temperature are generated. The model was capable of reproducing the coefficient of skewness for 50% of the records at the four locations tested. The lag one monthly coefficient of correlation was well reproduced for 65% of the records.

Given that the historic weather records are a sample of the historic weather, there will always be some kind of operational bias when using any particular probability distribution to generate synthetic weather records. The error associated with the estimation of the distribution parameters increases with the order of the moments. That is, the error associated with estimating the variance is worse than that associated with the mean and the error associated with the coefficient of skewness is worse than that associated with the variance.

One technique that can be used to minimize this operational bias is regionalization (i.e., developing a model for each geographical region). The problem with regionalization is that it destroys the intent of developing a general monthly precipitation and temperature model for all locations. It is believed that for operational purposes, developing regional models will be so cumbersome that it will be impractical. For the preliminary design of water resources works and wastewater land disposal sites, a weather model such as the one presented is appropriate for monthly generation of precipitation and temperature.

RESUMEN

Se evaluó un modelo matemático que puede usarse para producir series de precipitación y temperatura mensual que se asemejen a los datos históricos en términos de promedio, varianza, asimetría y correlación mensual rezagada. El esquema básico usado es el modelo de Markov de primer orden con una variable al azar con distribución gamma.

El modelo se evaluó en tres localidades en los Estados Unidos continentales (Omaha, Nebraska; Ithaca, Nueva York y San Francisco, California) y una localidad en Puerto Rico (San Juan). Los resultados de la evaluación indican que el modelo reproduce los primeros dos momentos de la distribución de datos históricos, promedio y varianza, de precipitación y temperatura muy bien. El modelo reprodujo el 50 por ciento de los coeficientes de asimetría de los datos históricos en las cuatro localidades en las que se evaluó. El coeficiente de correlación mensual fue bien reproducido para el 65 por ciento de los datos.

A pesar de que este trabajo fue motivado por la necesidad de series de datos sintéticos de precipitación y temperatura para el diseño preliminar de sistemas de aplicación de desperdicios a la tierra, su uso no se limita a ese tipo de proyecto.

LITERATURE CITED

1. Dávila, R. F., 1984. General Methodology for the Optimal Preliminary Design of Land Treatment Systems of Wastewater. Unpublished PhD Thesis. Cornell University, Ithaca, New York. 256 pp.
2. Loucks, D. P., J. R. Stedinger and D. A. Haith, 1981. Water Resource Planning and Analysis. Prentice-Hall Inc., Englewood Cliffs, N.J.
3. Matalas, N. C., 1967. Mathematical Assessment of Synthetic Hydrology, Water Resources Research, 3 (4): 932–45.
4. Selvalingam, S. and M. Miura, 1978. Stochastic Modelling of Monthly and Daily Rainfall Sequences. Water Resources Bulletin, 14 (5): 1105–119.
5. Snedecor, G. W. and W. G. Cochran, 1980. Statistical Methods, 7th ed, The Iowa State University Press.
6. Steel, R. G. D. and J. H. Torrie, 1980. Principles and Procedures of Statistics: A *Biometrical Approach*, 2nd ed, McGraw-Hill Book Co.
7. Thomas, H. A. and M. B. Fiering, 1962. Mathematical Synthesis of Streamflow Sequences for the Analysis of River Basims by Simulation, pp. 459–93. Harvard University Press.
8. U.S. Environmental Protection Agency, 1981. Process Design Manual—Land Treatment of Municipal Wastewater. Rep. No. EPA-625-1-81-013, U.S. Environmental Protection Agency, Cincinnati, Ohio.