Journal of Hydrology (NZ) 39(1): 65-82, 2000

A procedure for generating daily rainfall and evaporation data: an evaluation and some applications

D.R. Scotter[1], D.J. Horne[1] and S.R. Green[2]
[1]*Institute of Natural Resources, Massey University, Palmerston North, New Zealand.*
[2]*Environment and Risk Management Group, HortResearch, Palmerston North, New Zealand*

Abstract

An existing procedure for generating realistic daily rainfall data from the monthly average rainfall and number of wet days is tested and modified for New Zealand. Both the frequency distribution of extreme daily events at Palmerston North, and the 10 and 90 percentile monthly rainfall at eight contrasting locations, are accurately simulated. However east coast sites needed different treatment to other locations. Daily values for the Priestley-Taylor reference crop evaporation are also generated using average monthly sunshine hours and temperature values. The observed 33% lower evaporation on wet days (≥ 1 mm rainfall) than dry days is taken into account.

The generated values are used to construct a daily soil water balance for a soil near Palmerston North, and to obtain the probability density functions for the annual irrigation requirement, the annual amount of runoff, and the date when the soil is first dry enough to plough in spring. These probability density functions are similar to those obtained using 25 years of daily weather data.

If the effects of climate change, or El Niño-Southern Oscillation phenomena, on the monthly rainfall statistics at a location can be estimated, the approach offers a probabilistic way of looking at the implications of such factors for soil water management.

Introduction

Daily soil water balance models, run for extensive periods, have a range of uses, including planning for irrigation and effluent treatment (Green *et al.* 1999), and extrapolating the results of agronomic field trials. Such models however need daily values of rainfall and temperature and sunshine hours, which are not always readily available. In contrast, summaries of monthly averages for these variables have been published for many New Zealand locations (e.g. New Zealand Meteorological Service, 1983), or can be estimated by comparing data measured over a relatively short period with data from nearby sites with longer records. Thus it would be useful if daily soil water balances could be calculated using realistic sets of simulated daily data, generated from these average monthly values.

Of the data needed, rainfall data are the most difficult to generate, as they have the most day-to-day variability and the most skewed distribution. As Wilks and Wilby (1999) point out in a recent review, seminal ideas on how to do this entered the literature in 1962 and 1975. A Markov chain can be used to generate the distribution of wet days, and an approximately exponential distribution used to describe the amount of rainfall on those wet days. A number of schemes using these ideas have since been published, allowing synthetic daily rainfall data to be generated from monthly averages of rainfall and the number of wet days (e.g. Geng *et al.*, 1986). However, the validity of such schemes for New Zealand does not seem to have been checked. The approach has been calibrated and verified for a number of continental European and North American sites, but it appears not to have been tested in New Zealand. In addition, in its climate summaries the New Zealand Meteorological Service uses 1mm as the lower limit for defining a wet day. In contrast, Geng *et al.* (1986) imply that a wet day is one with measurable rainfall, and this difference needs to be taken into account.

As well as rainfall data, values for the reference crop evaporation (E_r), or potential evapotranspiration, are required to calculate the soil water balance. As defined by Penman (1956), this is the evaporation from an extensive area of well-watered, full-cover, short green vegetation. To a good approximation, it is independent of soil and vegetation type, and dependent only on the weather. For any month, or day of that month, E_r is a lot less variable than rainfall. But when daily values for E_r are generated from average monthly values, the negative correlation between wet days and the daily reference-crop evaporation needs to be considered. Cloudy days are associated with both rain and lower E_r values.

This paper describes how synthetic daily rainfall and reference-crop evaporation data may be generated from the average monthly values of four readily available variables, and considers the accuracy of the generated data. Then it gives examples of the use of such data to compute a soil water balance, and of the use of that water balance in soil water management.

Theory

Microsoft Excel has been used for computations. To assist the reader, some indications as to how this has been done are included.

Rainfall

The approach used here is essentially that employed in ClimGen, a programme for generating climate data (Campbell, 1990). Weather systems tend to produce rainfall over several days. Thus wet and dry days tend to be clumped together, and this needs to be taken into account. Generating daily rainfall involves two steps. The first step determines whether or not the day is a rain day. This is decided using a two-state, first-order Markov chain, as described by Geng *et al.* (1986). For each day, a random number between 0 and 1 is generated, using the worksheet function RAND(). Whether or not the day is wet is determined by comparing (using an IF function) the generated random number with one of two probabilities. One probability, $P(W|D)$, applies if the preceding day is a dry one; the other probability, $P(W|W)$, applies if the preceding day is rainy. $P(W|D)$ is of course lower than $P(W|W)$. Geng *et al.* (1986) suggest that

$$P(W|D) = bf \tag{1}$$

where f is the average fraction of days that are wet in the month, and b is a constant that they found empirically to have a value of about 0.75 for sites in the USA, the Philippines and the Netherlands. One aim of this study was to see if their b value of 0.75 is valid for a New Zealand site. The overall probability of a day being rainy, f, must equal the sum of the products of the two conditional probabilities and the probability of either a dry or wet day, that is

$$f = bf(1-f) + P(W|W)f. \tag{2}$$

Rearrangement gives

$$P(W|W) = 1 - b + bf. \tag{3}$$

Figure 1 shows $P(W|D)$ and $P(W|W)$ as a function of f, assuming $b = 0.75$. Regardless of the average fraction of wet days, the probability of rain on the day following a wet day is 0.25 greater than the probability of rain following a dry day.

If the day is a dry day, no second step is needed. For wet days, a second step generates the amount of rainfall. Rather than use a gamma distribution to do this, as in Geng *et al.* (1986), we follow Selker and Haith (1990) and

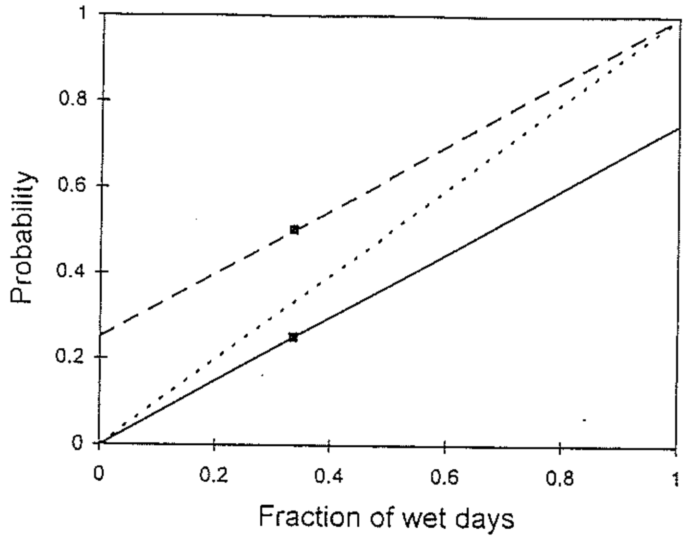

Figure 1 – Transitional probabilities for a wet day following a dry day (———), and a wet day following a wet day (– – – –), as a function of the fraction of wet days. Also shown is the 1:1 line (·······), which is the overall probability that a day will be wet. The two squares are the $P(W|D)$ and $P(W|W)$ values for Palmerston North found using the 25 years of data as a single set.

use a simpler Weibull-type distribution. If $F(x)$ is the cumulative probability distribution for wet day precipitation x, they suggest that

$$F(x) = 1 - \exp\{-[\Gamma(1 + 1/c)x/x_m]^c\} \tag{4}$$

where c is a dimensionless parameter, Γ is the gamma function, and x_m is the mean or expected value of x.

As mentioned above, the New Zealand Meteorological Service defines a wet day as having at least 1 mm of rain. Values of x_m, found by dividing the average monthly rainfall by the average number of wet days, thus assume that no wet days have less than 1 mm of rain. To take account of this, and to still get the correct total monthly amount of rain on average, equation (4) needs to be modified to

$$F(x) = 1 - \exp\{-[\Gamma(1 + 1/c)(x - 1)/(x_m - 1]^c\} \tag{5}$$

Solving for x gives

$$x = [(x_m - 1)/\Gamma(1+1/c)][1 - \ln(1-F)]^{1/c} + 1 \quad (6)$$

From an analysis of United States rainfall data, Selker and Haith (1990) suggested 0.75 as a suitable value for c. Equation (6) then simplifies to

$$x = [(x_m - 1)/1.191][-\ln(1-F)]^{1.33} + 1 \quad (7)$$

The other value of c that is considered is 0.5. With this value, equation (6) becomes

$$x = [(x_m - 1)/2][-\ln(1-F)]^2 + 1 \quad (8)$$

Figure 2 shows F as a function $(x-1)/(x_m-1)$ for c values of 0.75 and 0.5.

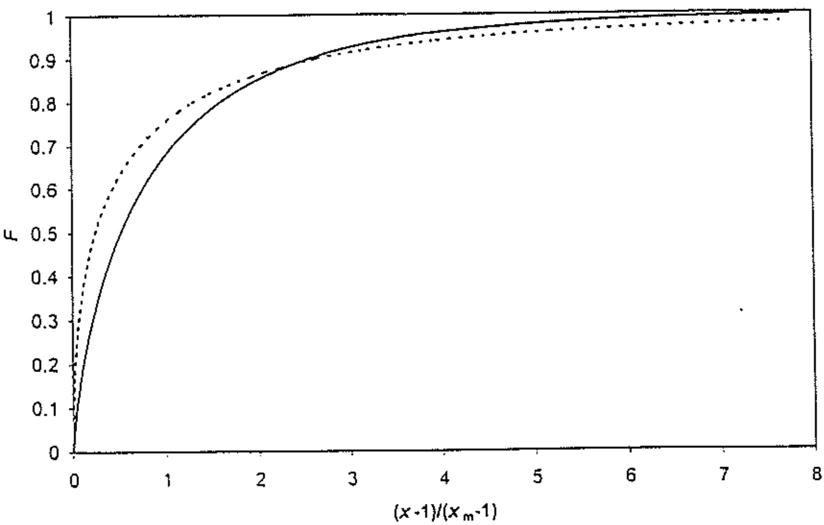

Figure 2 - The cumulative probability distribution of $(x-1)/(x_m-1)$ for c values of 0.75 (———) and 0.5 (··········).

The lower c value generates more extreme rainfall events. To use equations (6), (7) or (8), another random number between 0 and 1 is generated for each wet day, and put equal to F. The equations then give a value for x, the rainfall for the day.

Reference-crop evaporation

The equation of Priestley and Taylor (1972) is used to compute the reference crop evaporation (E_r). McNaughton *et al.* (1979) showed that, for well-watered pasture and full-cover crops growing near Palmerston North, it was as accurate as the Penman equation, which requires more data. First, average monthly values for the reference crop evaporation are found for the site of interest, or a location not too far away. The data needed are the latitude, average monthly values for the screen temperature, and the number of sunshine hours. This is done in the following manner. The Priestley-Taylor equation is

$$E_r = \frac{1.26 s R_n}{\rho_w L(s+\gamma)} \tag{9}$$

where s is the slope of the relationship between the saturated vapour density and temperature, γ is the psychrometric constant, R_n is the net radiation over the 24 hour period (MJ m^{-2}), ρ_w is the density of water (1000 kg m^{-3}), and L is the latent heat of vaporization of water at ambient temperature (2.5 MJ kg^{-1}). The soil heat flux density is not included in equation (9), as over our 24 h periods it is small relative to the net radiation (McNaughton *et al.*, 1979). A quadratic equation fitted to tabulated values for the dimensionless ratio of $s/(s + \gamma)$ at an air pressure of 100 kPa over the temperature range 5 to 20°C (from Tanner, pers. com.) gives

$$s/(s+\gamma) = 0.403 + 0.0164 T_{av} - 0.00012 T_{av}^2 \tag{10}$$

where T_{av} (°C) is the average screen air temperature for the month. The net radiation may be estimated from the daily incoming solar radiation (R_s) (Scotter *et al.*, 1979) as

$$R_n = 0.62 R_s - 1.47, \tag{11}$$

where both R_n and R_s are in units of MJ m^{-2}. The average daily solar radiation is estimated from the average number of sunshine hours per day for the month as follows (note all angles are in radians).

First the solar declination angle, δ, is approximated (Rosenberg 1974) as

$$\delta = 0.41 \sin[2\pi(M - 80)/365] \tag{12}$$

where M is the Julian day for the midpoint of the month. The half-day length, H, expressed as an angle, is next found as

$$H = \cos^{-1}(-\tan\varphi \tan\delta) \tag{13}$$

where φ is the latitude. The solar radiation received in a day (in MJ m^{-2}) by a horizontal surface outside the earth's atmosphere (R_o) is found (Sellers, 1965) as

$$R_o = 38.5(H \sin\varphi \sin\delta + \cos\varphi \sin H)(d_{av} / d)^2 \tag{14}$$

where d is the distance of the earth from the sun on the day in question, and d_{av} is the average distance. Fitting values from Robinson (1966) to a sine wave gives

$$d_{av} / d)^2 = 1 + 0.0334 \sin(0.0172M - 4.81). \tag{15}$$

Note that $(d_{av}/d)^2$ is always within 3.5% of unity. Next, the maximum number of sunshine hours measurable on the mid-month day (N) is found as

$$N = 24H / \pi - 0.5. \tag{16}$$

N is the actual maximum, less half an hour, due to the inability of the Campbell-Stokes sunshine recorder to register at low solar elevations (de Lisle, 1966). Lastly the incoming solar radiation for a flat site is estimated as

$$R_s = R_o(0.25 + 0.54n / N), \tag{17}$$

where n is the average number of sunshine hours measured (de Lisle, 1966). Once R_s has been estimated, then the average daily reference crop evaporation for the month can be estimated from equations (11) and (9).

A numerical example of the use of the above equations may help the reader. For September in Palmerston North (E05363), N. Z. Meteorological Service (1983) gives the average screen temperature as 10.6°C, and the average number of sunshine hours measured as 133. Equation (10) gives $s/(s+\gamma)$ as 0.56. The middle of the month, 15 September, is Julian day 258 (except in a leap year), so equation (12) gives the declination angle as +0.025. The latitude is 40°23' S or –0.7048 radians, so equation (13) gives H as 1.55. Equation (15) gives $(d_{av}/d)^2$ as 0.99, and equation (14) gives R_o as 28.1 MJ m^{-2}. Equation (16) gives N as 11.3 h. As n is 133/30 or 4.4 h, equation (17) gives R_s as 13.0 MJ m^{-2}. Equation (11) then gives R_n as 6.56 MJ m^{-2}, and lastly equation (9) gives the average September value of E_r as 1.9 mm/day.

The actual daily values are assumed to be normally distributed around the mean monthly values. However, the expected negative correlation between wet days and solar radiation, and therfore E_r, needs to be taken into account.

To do this the following approach is adopted. Define a parameter a such that

$$\langle E_{\mathrm{d}} \rangle = a \langle E_{\mathrm{w}} \rangle \tag{18}$$

where $<E_d>$ and $<E_w>$ are the average values of E_r for dry and wet days in a certain month. The average daily reference crop evaporation for the month, $<E_r>$, is related to the other two averages by

$$\langle E_{\mathrm{r}} \rangle = \frac{j\langle E_{\mathrm{w}} \rangle + (J - j)\langle E_{\mathrm{d}} \rangle}{J} \tag{19}$$

where J is the number of days in the month and j is the average number of wet days in the month. Combining and rearranging equations (18) and (19) gives

$$\langle \mathrm{E_w} \rangle = \frac{\langle E_{\mathrm{r}} \rangle}{a + (1 - a)\, j/J}. \tag{20}$$

To generate a value for E_r, an IF function selects either $<E_w>$ or $<E_d>$ as the appropriate mean value, based on whether the rainfall already generated is zero or not. A constant value for C, the coefficient of variation, is assumed for E_w and E_d, and for all months. This assumption is justified below. The appropriate standard deviation is estimated as C multiplied by the mean. A random number between 0 and 1 is then generated, and used as the cumulative probability in the inverse normal distribution worksheet function (NORMINV), along with the appropriate mean and standard deviation, to generate a value for E_r on that day.

The soil water balance

The water balance model of Scotter *et al.* (1979) was used. They describe a relationship between the soil water deficit and the actual evaporation, derived from measurements under pasture growing on Tokomaru silt loam near Palmerston North. It assumes that the soil water deficit (relative to field capacity) needs to be greater than about 100 mm before the actual evaporation rate drops below the reference crop rate. Also, a topsoil available water store, holding up to 25 mm, is included. This can supply water for evaporation at the reference crop rate following rain, over-riding any effect of the total soil water deficit. Runoff (as deep percolation, surface runoff and/or drain flow) is assumed to occur only when there is rainfall in excess of that needed to fully recharge the soil profile to field capacity. Interception is not treated explicitly. For pasture it is low, and usually just replaces an equivalent amount of transpiration, due to the boundary-layer resistance being greater than the canopy resistance (McNaughton and Jarvis, 1983).

Evaluating parameters and testing theory

Rainfall

The rainfall generation algorithms were evaluated using 25 years of data for Palmerston North, covering the period from 1974 to 1999. Over the 9132 days there were 3062 days with ≥ 1 mm of rain, giving an average f value of 0.335, and the two conditional probability values for wet days following dry or wet days shown in Figure 1. The overall value found for b was 0.748, the same as the average value of 0.75 found by Geng *et al.* (1986). For individual months in Palmerston North, the highest and lowest f values were 0.42 (for July) and 0.24 (for February). The b values found for these two months were 0.74 and 0.73. Similarly, for Wellington, Tomlinson (1992) comments that on average there are 125 rain days per year, so the fraction

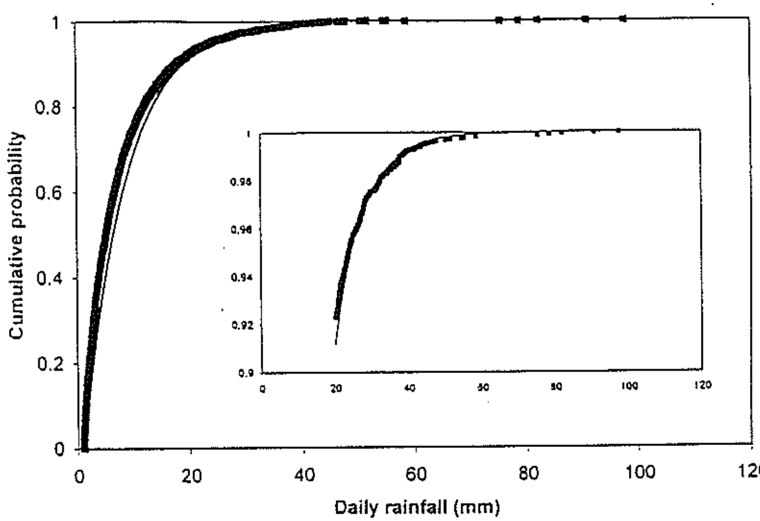

Figure 3 – The cumulative distribution of daily rainfall events ≥ 1 mm for Palmerston North for the period from June 1974 to May 1999 (■), and the distribution given by equation (7) (———). The inset shows an expanded view of the observed (■) and generated (———) cumulative probability distributions for rainfall events > 20 mm.

of wet days, and overall probability of a wet day, is 0.34. Then he gives the probability of a wet day if the previous day is a rain day as about 0.5, which from equation (3) implies b equals 0.76. These results show that the assumption $b = 0.75$ is valid for Palmerston North and Wellington. Further testing is needed for other New Zealand locations.

Next the validity of equation (7) was tested, again using 25 years of rainfall data for Palmerston North. N.Z. Meteorological Service (1983) shows that x_m, the average rainfall on a wet day, is relatively constant throughout the year, varying from 7.1 mm in November to 9.6 mm in February (with an overall average of 7.99 mm), so the data were treated as a single population. The observed cumulative probability distribution of wet-day rainfalls is shown in Figure 3, along with the distribution generated using equation (7). For 90% of wet days, which get less than 20 mm of rain, the predicted amount of rain is slightly higher than that observed. This is due to the rainfalls of less than 1 mm/day on 'dry' days (on average 21 mm/year). To get the correct average monthly rainfall, equation (5) distributes this rainfall as extra rainfall on the wet days. What is important, however, is that the probability of large daily events is correctly described. The inset in Figure 3 shows an expanded view for rainfalls greater than 20 mm. For these events the observed and generated probabilities are in close agreement.

Next, the approach was tested further using data from eight locations, covering Northland to Otago, East Coast to West Coast, and high to low rainfall. A particular month was selected for each site so as to cover the most extreme behaviours, in terms of high and low rainfall, and the 'gamma parameter', which indicates the skewness of the distribution of rainfall for the month (N.Z. Meteorological Service, 1979). The lower the value, the more skewed the distribution. Table 1 shows the average monthly rainfall, the average number of rain days (with ≥ 1.0 mm of rain), and the 90 percentile and 10 percentile rainfall values, all taken from N.Z. Meteorological Service (1983). Only sites for which at least 50 years of data had been analysed to obtain these statistics were chosen. The selected values for average monthly rainfall ranged from 18 mm to 278 mm, and the mean number of rain days per month from 4 to 17. Gamma parameter values varied from 1.4 for Hastings in January to 7.6 for Ruakura in August.

For each of the month and location combinations selected, 100 years of rainfall data were generated using equations (1), (3), (7) and (8), and the 10 and 90 percentiles of these 100 monthly rainfall values were found. This was done 10 times, and the mean and standard deviation of the ten 10 and 90 percentile values calculated. The observed and generated means were very close, as expected, and are not shown. The 10 and 90 percentile values are shown in Table 1. Except for the three east coast locations, all the observed values fell within one standard deviation of the mean of the generated values

Table 1 – Observed and generated rainfall statistics for 8 locations and selected months. All rainfall values are in mm/month. Means and standard deviations are given for the generated 10 and 90 percentile values. Other details are given in the text.

Location	Waipoua	East Cape	Ruakura	Hastings	Palm. North	Hokitika	Christchurch	Alexandra
Latitude (S)	35 39	37 42	37 47	39 39	40 23	42 43	43 32	45 16
Long (E)	173 33	178 33	175 19	176 51	175 37	170 57	172 37	169 23
Month	October	January	August	January	July	October	December	August
				Measured Values				
Rain days	15	8	14	7	12	17	7	4
Mean rain	139	90	110	65	89	278	58	18
90%	210	182	175	136	140	395	119	35
10%	73	18	60	13	38	160	13	3
				Generated Values				
90%(c=0.75)	216 ±15	160 ± 9	169 ±11	114 ± 9	137 ± 5	414 ±21	103 ± 7	37 ± 3
90%(c=0.5)	232 ±19	178 ±24	189 ±21	133 ±18	154 ±16	472 ±33	117 ±10	42 ± 6
10%(c=0.75)	75 ± 6	32 ± 2	57 ± 6	19 ± 4	42 ± 4	160 ±18	19 ± 3	2 ± 1
10%(c=0.5)	59 ± 5	21 ± 4	43 ± 5	14 ± 3	34 ± 4	112 ±10	13 ± 3	2 ± 1

when equation (7) with c of 0.75 was used. For the three east coast locations (East Cape, Hastings and Christchurch) the difference is greater than one standard deviation. But when equation (8), with $c = 0.5$, is used for these locations, the observed 90 percentile and 10 percentile values fall within one standard deviation of the mean of the generated values for these locations. The distinctive behaviour of the three east coast locations is consistent with the coefficient of variation of the mean annual rainfall being higher there than in other areas (Tomlinson, 1992).

In view of the above, we suggest two approaches to generating daily rainfall. The simpler approach is to use equation (7) ($c = 0.75$) for any New Zealand location of agricultural interest except east coast locations. For east coast locations, we suggest the use of equation (8) ($c = 0.5$). The second and more complex, but more accurate, approach would be to use equation (6), with a c value optimised for the location and perhaps month of interest, using 20 or more years of actual rainfall data, as we have done in Figure 3 for Palmerston North. For most applications, the extra effort required for the second approach will probably not be warranted.

Reference-crop evaporation

Twenty-five years (1974-1999) of daily Priestley -Taylor values for Palmerston North were computed using actual data for screen temperatures and sunshine hours. Statistics for five months are given in Table 2. The monthly averages are all within 0.1 mm of values calculated as described above, using the monthly means of screen temperature and sunshine hours

Table 2 – Average daily values of reference crop evaporation (E_r) and values for *a* (the ratio of dry day to wet day reference evaporation) for five months in Palmerston North, computed from 25 years of weather data. Also shown are values for the coefficient of variation, *C*, for wet and dry day evaporation.

Month	January	March	July	October	December
E_r (mm)	4.2	2.7	0.7	2.7	3.9
a	*1.4*	*1.5*	*1.7*	*1.5*	*1.5*
C (wet days)	0.38	0.45	0.68	0.41	0.36
C (dry days)	0.31	0.36	0.52	0.36	0.32

for the 1935 to 1980 period (N.Z. Meteorological Service, 1983). Also, for all twelve months of the year, the average Priestley-Taylor values were within 0.1 mm/d of average Penman values for Palmerston North supplied by the New Zealand Meteorological Service (N.Z. Meteorological Service, pers. comm.), and the annual totals differed by only 6 mm, i.e. 845 mm for the Priestley-Taylor calculations performed as described above using average monthly data, and 839 mm for the Penman data supplied.

The reference crop evaporation values for dry days in Table 2 were on average 1.5 times greater than the wet day values (reflecting the different amounts of net radiation), so this value was used for a. The values for the parameter C, the coefficient of variation, were not considered different enough to warrant the use of different values for different months, or for wet and dry days. A value of 0.39 was used, the weighted average of the C values in Table 2. When computing this average, the two high values for July were omitted, due to the low, and so less important, evaporation then. Taking account of the negative correlation between evaporation and rain days increases evaporation in months with fewer than average wet days, but does not change the mean monthly evaporation. In the next section we comment on how this can affect irrigation requirements.

Practical applications based on the soil water balance

Water balance calculations can be used to approach a number of practical problems – three such problems are considered here. For each problem, calculations were made using both 25 years of actual meteorological data for Palmerston North, and using 1000 years of generated data. These data were generated in a few minutes, using a Visual Basic macro within Excel.

The first illustrative problem relates to irrigation design. How much water is needed to irrigate pasture around Palmerston North? The scheduling

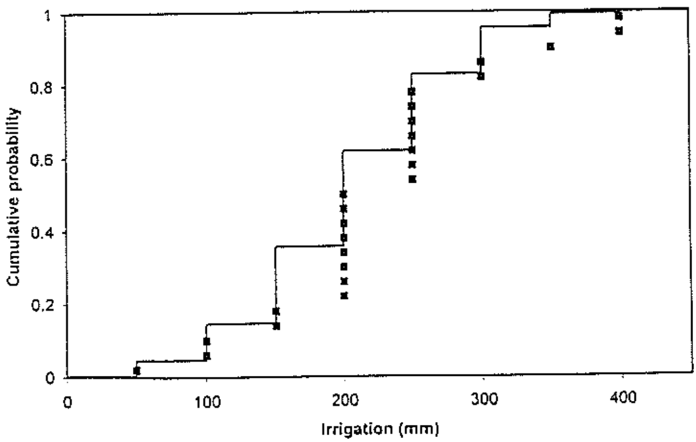

Figure 4 – The cumulative probability distribution of annual irrigation required, calculated using 25 years of actual weather data (▪) and 1000 years of generated data (———), and assuming 50 mm of irrigation is applied whenever the soil water deficit reaches 75 mm.

criterion used is that 50 mm of irrigation is added on the day following any day that the soil water deficit exceeds 75 mm. The distribution of the annual amount of irrigation is shown in Figure 4. The amount applied would vary from 50 to 400 mm, with 200 mm being the mode or most common value. The distributions computed using 25 years of actual weather data, and using 1000 years of generated data are very similar, and a two-sided t-test shows that the means are not significantly different at the 90% confidence level. Data such as these would be useful for a person applying for a Water Right to irrigate, or for an organization administering such Water Rights.

In developing and running the model, some effort was made to account for the negative correlation between evaporation and wet days. To assess the importance of this effort, the simulation was run again, but without adjusting for this correlation (i.e. with a = 1). The average irrigation requirement decreased from 202 mm to 195 mm (significantly different at 1%), and the standard deviation of the values fell from 75 mm to 71 mm. So while the effect of the correlation is not large, it is discernible. As would be expected, it tends to increase the predicted need for irrigation, and to make it more variable.

The second illustrative problem relates to water harvesting. How much surface runoff and drain flow occurs over the winter-spring period around Palmerston North from a soil like Tokomaru silt loam? Heng *et al.* (1991)

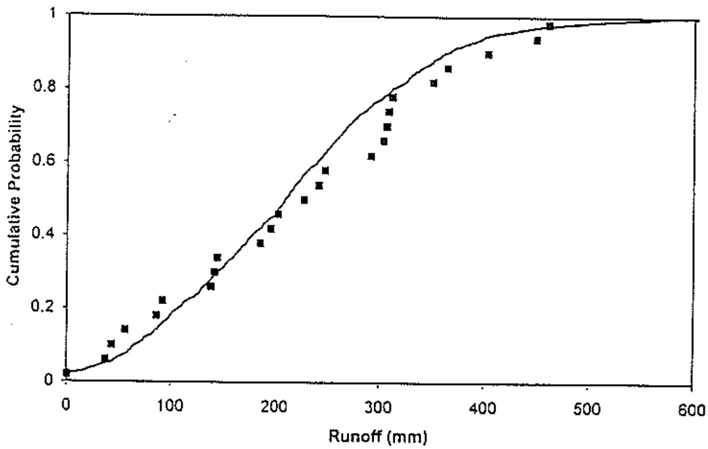

Figure 5 – The cumulative probability distribution of annual runoff, calculated using 25 years of actual weather data (▪) and 1000 years of generated data (———).

demonstrated that very little deep percolation occurs within this soil, so nearly all runoff from mole-pipe drained land appears as either pipe drainage or surface runoff. Therefore runoff values from the model can be used to indicate the likelihood that a water-harvesting dam in an ephemeral water-course will fill over the winter-spring period (Turner *et al.*, 1977). Figure 5 shows the results. Again there is close agreement between the distributions obtained using actual and generated meteorological data; the probability that the means are different is less than 50%. Of course, a much smoother probability curve is obtained with 1000 years of generated data than with just 25 years of actual data. The average yearly runoff is about 220 mm, but note the huge range from 0 to over 400 mm. In about 20% of years, less than 100 mm of runoff occurs from unirrigated pasture. That is less than 1000 m^3 of water per hectare. In 2% of the generated years, and one of the 25 actual years, no runoff at all is expected.

The third problem is: when is the soil dry enough in spring either to cultivate, or to apply 20 mm of dairy shed-effluent without causing immediate runoff? To model this, the first day after 31 July when the soil water deficit was greater than 20 mm is extracted from the water balance. The result is shown in Figure 6. In this case, two distributions obtained using generated data are shown. One was found using the long-term average (1928-1980) monthly rainfall and number of wet days, and the other was found using the average monthly rainfall and number of wet days for the last 25 years (1974-

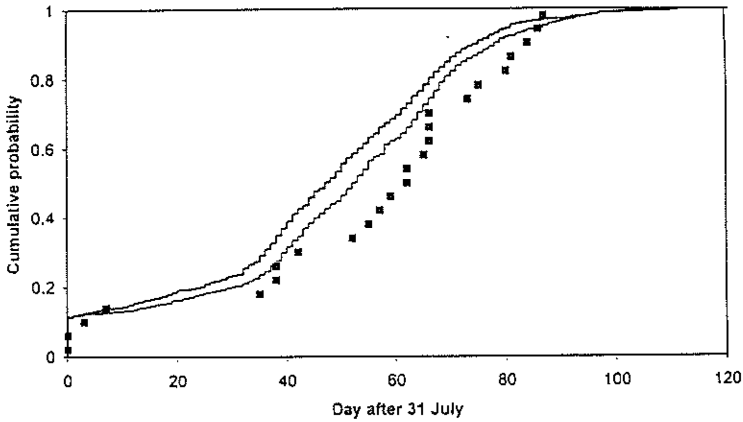

Figure 6 - The cumulative probability distribution of the number of days after 31 July when the soil water deficit is greater than 20 mm, calculated using 25 years of actual weather data (■) and 1000 years of generated data using long term averages (heavy line), and averages for the 25 years (lighter line).

1999). The data points in the third distribution shown were obtained using the 25 years of actual weather data. All three distributions are similar, and the means are not significantly different at the 90% confidence level. However, the distribution generated using averages for the most recent 25 years is closer to the distribution obtained using the actual data for that period. The starting dates found for cultivation or effluent application range from 1 August to late October or early November.

This third application of the model tests it more stringently than the other two. When the differences between daily rainfall and evaporation are integrated over several seasons to obtain the annual runoff or the irrigation required, considerable statistical averaging occurs. While the timing and size of individual rainfall events is important, it is not critical. In contrast, in the third application the model must be able to provide accurate descriptions of the rainfall and evaporation distributions over periods of approximately a week, to successfully identify the date when the soil water deficit first reaches 20 mm and spring field cultivation can begin. Figure 6 thus provides the strongest evidence of the robustness of the model.

The rainfall/evaporation model and climate change

Our approach assumes that the weather has a 'memory' of one day, i.e. that the rainfall and evaporation on a certain day are affected by whether or not it has rained on the preceding day. Other than that, weather events are assumed to be random in nature. But it is well known that climate changes over geological time scales, and that due to global warming, changes over much shorter time scales are likely. We also know that from year to year there are climatic patterns. So, what about climate change?

As our approach assumes that the statistical parameters for rainfall and the reference crop evaporation are stationary, the short answer is that such effects are ignored. However, if techniques were available to predict changes in the mean monthly rainfall and/or number of rain days at a location, the above approach could be used to predict the effects of such a change simply by varying x_m. For example, climatologists have found a correlation between spring rainfall over north-eastern New Zealand and the Southern Oscillation Index (SOI) (Gordon, 1986; Mullan, 1996). The approach described here could be used to turn predictions made using such relationships into a form that is of direct use to soil water managers.

Conclusions

The rainfall generation approach works well at eight very different New Zealand locations, although east coast locations require slightly different treatment. The procedure to generate daily values for the reference crop evaporation takes into account the observed 33% lower evaporation on wet days (≥ 1 mm rainfall) than dry days.

One thousand years of generated data were used to obtain the probability density functions of irrigation requirements, annual runoff, and earliest spring cultivation dates. The probability density functions obtained using the generated weather data were not significantly different to ones obtained using 25 years of actual weather data for Palmerston North.

If the effects of climate change, or El Niño-Southern Oscillation (ENSO) phenomena, on the monthly rainfall, temperature and sunshine hour statistics at a location can be estimated, the approach offers a probabilistic way of predicting the effects of climate change in a form of interest to soil water managers.

Acknowledgement

The constructive criticism of three reviewers and an editor is acknowledged.

References

Campbell, G.S. 1990: ClimGen, climatic data generator. Washington State University. Available at http://www.bsyse.wsu.edu/climgen (verified 26 Oct, 1999)

de Lisle, J.F. 1966: Mean daily insolation in New Zealand. *New Zealand Journal of Science 9*: 992-1005.

Geng, S.; Penning de Vries, F.W.T.; Supit, I. 1986: A simple method for generating daily rainfall data. *Agricultural and Forest Meteorology 36*: 363-376.

Gordon, N.D. 1986: The Southern Oscillation and New Zealand weather. *Monthly Weather Review 114*: 371-387.

Green, S.R.; Clothier, B.E.; Mills, T.M.; Millar, A. 1999: Risk assessment of the irrigation requirements of field crops in a maritime climate. *Journal of Crop Production 2*: 353-377.

Heng, L.K.; White, R.E.; Bolan, N.S.; Scotter, D.R. 1991: Leaching losses of major nutrients from a mole-drained soil. *New Zealand Journal of Agricultural Research 34*: 325-334.

McNaughton, K.G.; Clothier, B.E.; Kerr, J.P. 1979: Evaporation from land surfaces. *In*: Murray, D.L.; Ackroyd, P. *(eds.) Physical hydrology: New Zealand experience.* New Zealand Hydrological Society, Wellington, 97-119.

McNaughton, K.G.; Jarvis, P.G. 1983: Predicting effects of vegetation changes on transpiration and evaporation. *In*: Kozlowski, T.T. *(ed.) Water Deficits and Plant Growth Vol VII.* Academic Press, New York, 1-47.

Mullan, B. 1996: Effects of ENSO on New Zealand and the South Pacific. *In*: Braddock, D. *(ed.) Prospects and Needs for Climate Forecasting.* Miscellaneous Series 34, The Royal Society of New Zealand, Wellington.

New Zealand Meteorological Service 1979: Rainfall parameters for stations in New Zealand and the Pacific Islands. *N.Z. Meteorological Service Miscellaneous Publication 163.*

New Zealand Metorological Service 1983: Summaries of climatological observations to 1980. *N.Z. Meteorological Service Miscellaneous Publication 177.*

Penman, H.L. 1956: Evaporation; an introductory survey. *Netherlands Journal of Agricultural Science 4*: 9-29.

Priestley, C.H.B.; Taylor, R.J. 1972: On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review 100*: 81-92.

Robinson, N. 1966: *Solar Radiation.* Elsevier, Amsterdam.

Rosenberg, N.J. 1974: *Microclimate: The Biological Environment.* Wiley, New York.

Scotter, D.R.; Clothier, B.E.; Turner, M.A. 1979: The soil water balance in a Fragiaqualf and its effect on pasture growth in Central New Zealand. *Australian Journal of Soil Research 17*: 455-65.
Selker, J.S.; Haith, D.A. 1990: Development and testing of single-parameter precipitation distributions. *Water Resources Research 26*: 2733-2740.
Sellers, W.D. 1965: *Physical Climatology*. The University of Chicago Press, Chicago.
Tomlinson, A.I. 1992: Precipitation and atmosphere. *In*: Mosley, P. (*ed.*) *Waters of New Zealand*. N.Z. Hydrological Society, Wellington, 63-74
Turner, M.A.; Scotter, D.R.; Bowler, D.G.; Tillman, R.W. 1977. Water harvesting: the concept and use in a humid climate. Proceedings of Soil and Plant Water Symposium. DSIR Information Series Number 126.
Wilks, D.S.; Wilby, R.L. 1999. The weather generation game; a review of stochastic weather models. *Progress in Physical Geography 23*: 329-357.

Manuscript received: 26 November 1999; accepted for publication: 3 March 2000.