551.577.3 : 311.15

DAILY RAINFALL PATTERNS AT MELBOURNE

By G.A. Watterson and M.P.C. Legg

Department of Mathematics,

Monash University, Victoria

(Manuscript received March 1967)

ABSTRACT

Days during the 107 years 1859-1965 have been classified as either wet or dry with respect to the rainfall recorded at Melbourne. The paper analyses the hypothesis that the resulting pattern can be adequately fitted (if not explained) by a first-order Markov chain model, in which the probability of a day being dry depends on time of year and on the state of the previous day, but is not further influenced by the state of earlier days.

The model is shown to fit reasonably well the observed patterns of behaviour in three-day intervals, in spells of consecutive dry days, in spells of consecutive wet days, and in dry-wet cycles. The conclusions are in general agreement with the work of Gabriel and Neumann for Tel Aviv winter rainfall.

1. INTRODUCTION

There have been many studies of rainfall data, including investigations of periodicities, the effects of cloud-seeding, phase of the moon, day of the year, etc. The present investigation parallels the work of Gabriel and Neumann (1957, 1962), who studied Tel Aviv rainfall to see if the wet-, and dry-day pattern could be adequately described by a Markov chain model. In our study, we consider the hypothesis that Melbourne's daily rainfall can be described similarly.

The Bureau of Meteorology, Australia, has kindly provided us with daily rainfall data for Melbourne for the 107 years 1859-1965, on 2568 punched cards. There is much that could be done with such an array of data; we have limited ourselves to consider those aspects most closely related to a Markov chain hypothesis for the wet-, and dry-day pattern, namely the influence of previous days on the current day, the seasonal effect, the dry spell and wet spell lengths, and the cycle lengths.

It is thought that other aspects of Melbourne's rainfall might be analysed later with this hypothesis in mind; for instance, given that a day is wet, does the amount of precipitation depend on the previous day's precipitation? Can the Markov chain model be improved by incorporating extra parameters such as phases of the moon, sun-spots, and (for predictive purposes) prior rainfall at a different location, say Adelaide?

2. ANALYSIS OF THREE-DAY PATTERNS

The data were analysed on a C.D.C. 3200 computer; for each analysis the data were coded in the form 0 = no precipitation (dry), 1 = some precipitation (wet), where originally precipitation was recorded to the nearest 0.01 in. Thus, effectively, a dry day is one which has less than 0.005 in. precipitation; of course, other conventions could have been used.

For presentation of results, we will assume each year has 365 days numbered 1, 2, ..., 365; thus data for December 31 of a leap year will not be presented, although in some of our analyses we used such data as part of the history for days 1 and 2 of the succeeding year.

We consider first the question of whether the two states, 0, 1, can be considered to be the states of a first or a second order Markov chain. To this end, we classify every available triple of three consecutive days into one of eight types:

000, 001, 100, 101, 010, 011, 110, 111.

Denoting by $n_{ijk}(d)$ the number of years in which triples of type (i, j, k) occur on days (d - 2, d - 1, d) respectively, where i, j, k = 0 or 1, d = 1, 2, ..., 365, a sample of the computed results is given in Table 1.

Table 1. Frequencies of triple types

Day d	$n_{000}(d)$	$n_{001}(d)$	$n_{100}(d)$	$n_{101}(d)$	$n_{010}(d)$	$n_{011}(d)$	$n_{110}(d)$	$n_{111}(d)$	Total
1	50	10	11	1	14	6	7	7	106
2	45	16	17	4	5	6	9	4	106
3	51	12	12	2	6	14	3	7	107
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
189	13	10	10	12	10	17	11	24	107
190	13	10	11	10	8	14	13	28	107
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
364	44	13	11	4	11	7	12	5	107
365	37	18	17	6	9	8	5	7	107

The total column in Table 1 indicates that we have, generally, 107 years' data in all, but that for days 1 and 2 of a year we have 106 years of triples available (we did not have the records for the last two days of 1858 in order to include the triples overlapping 1858-1859).

From the triple numbers in Table 1, estimates of the conditional probability of a dry day on day d can be obtained, given the four types of two-day histories. Using the notation $p_{ij0}(d)$ for the probability that day d is dry given that day d - 2 is of type i (i = 0, 1) and day d - 1 is of type j (j = 0, 1), the estimates are

$$\hat{p}_{ij0}(d) = \frac{n_{ij0}(d)}{n_{ij0}(d) + n_{ij1}(d)}$$

The complementary quantities, $1 - \hat{p}_{ij0}(d)$, would be estimates of wet day probabilities under the same circumstances. Corresponding to Table 1, we have estimates as in Table 2.

By inspection of the complete listing corresponding to Table 2, it is clear that the day, d, does influence these dry-day probabilities which have a peak at (roughly) the end of January and a trough in mid-July. Moreover, the probability of getting a dry day is greater if the previous day was also dry. These remarks may not seem very profound to residents of Melbourne, of course! But our main concern is to investigate whether a first order Markov chain model, with day-dependent transition probabilities, gives an adequately good fit to the data, or whether account must also be taken of the type of day occurring two (or more) days earlier. If, for each d, $\hat{p}_{000}(d) \doteq \hat{p}_{100}(d)$ and $\hat{p}_{010}(d) \doteq \hat{p}_{110}(d)$ both hold within sampling fluctuations, the conclusion would be that the only significant influence day d - 2 exerts over day d, is via its influence on day d - 1, and a knowledge of the latter's type obviates any need to also consider d - 2's type.

Table 2. Estimated probabilities for a dry day

Day d	$\hat{p}_{000}(d)$	$\hat{p}_{100}(d)$	$\hat{p}_{010}(d)$	$\hat{p}_{110}(d)$
1	.8333	.9167	.7000	.5000
2	.7377	.8095	.4545	.6923
3	.8095	.8571	.3000	.3000
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
189	.5652	.4545	.3704	.3143
190	.5652	.5238	.3636	.3171
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
364	.7719	.7333	.6111	.7059
365	.6727	.7391	.5294	.4167

Although it would be possible to perform significance tests of the above hypothesis at each day separately, such tests would not be very powerful, as the probability estimates are based on relatively few observations and only large departures would likely be judged significant. We would still have the problem of deciding which of the 365 (or 2x365) tests appear significant by chance. Alternatively, if some (one-year-periodic) functions of d could be fitted to the four columns of Table 2, we could then get a reasonably powerful test for the equality or otherwise of the fitted functions, but no simple functions with the right sort of day-dependence come to mind. In Figure 1 we exhibit smooth versions of the transition probability estimates, obtained after taking 31-day weighted moving averages thus:

$$\hat{p}_{ij0}(d) = \sum_{m=-15}^{15} W_{ij}(d+m)\,\hat{p}_{ij0}(d+m) \Big/ \sum_{n=-15}^{15} W_{ij}(d+n) \qquad \ldots(1)$$

where the weights W_{ij} are the number of observations on which the estimates $\hat{p}_{ij0}$ were based, that is

$$W_{ij}(d+m) = n_{ij0}(d+m) + n_{ij1}(d+m), \; d = 1, 2, \ldots, 365 \qquad \ldots(2)$$

and where days are identified cyclically: 0 = 365, -1 = 364, -2 = 363, ... Of course, the smoothing greatly reduces the sampling variability, to a standard deviation around .02, but greatly increases the correlation between estimates for adjacent days.

It was finally decided to first test departures from hypothesis by computing weighted, linear, least-squares regressions for the four transition probabilities, within calendar months. Thus within each month, models of the form

$$\hat{p}_{ij0}(d) = a_{ij} + b_{ij}.d + \text{error}, \quad i, j = 0, 1 \qquad \ldots(3)$$

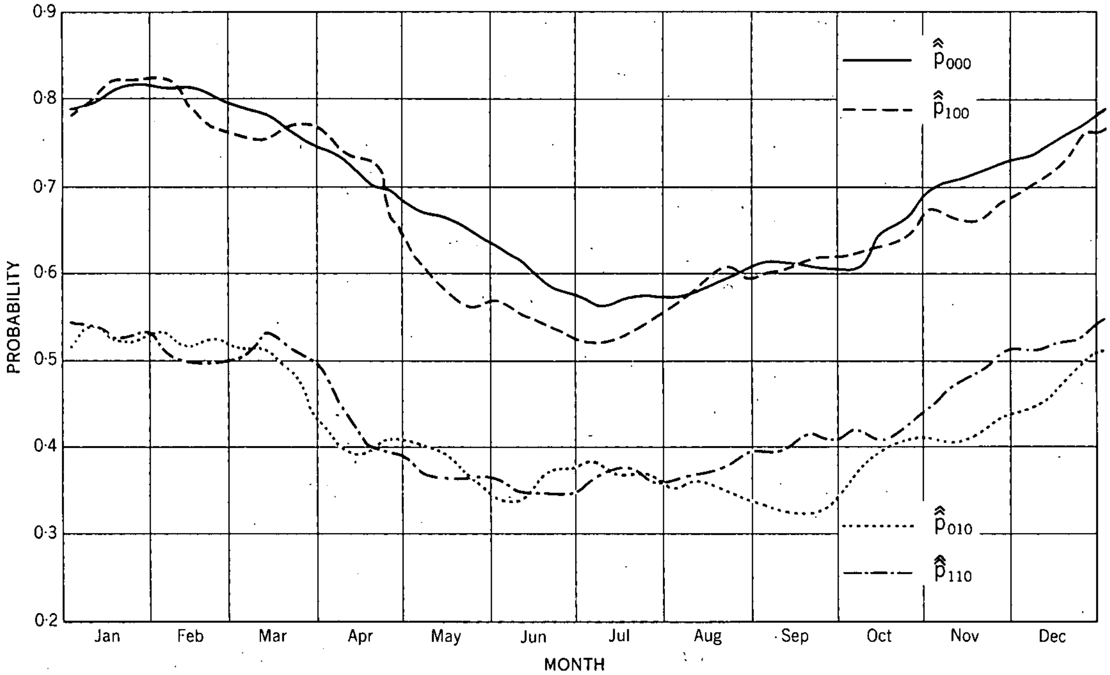

Fig. 1 Smoothed estimates of dry day probabilities, allowing for states of two previous days.

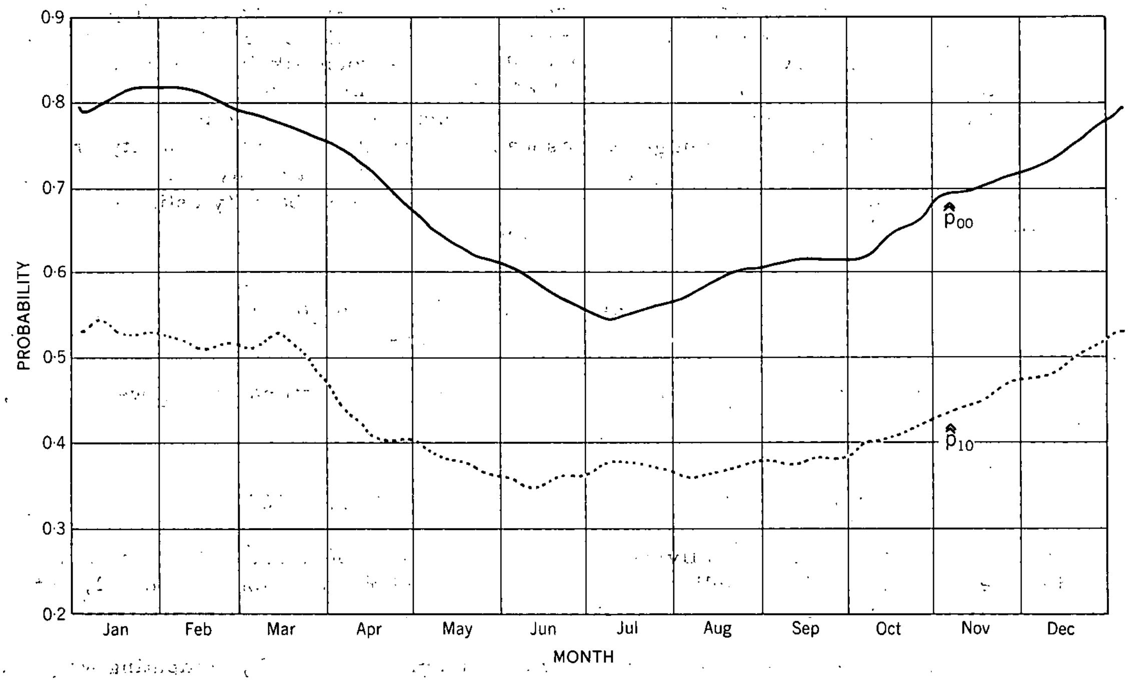

Fig. 2 Smoothed estimates of dry day probabilities, allowing for state of previous day.

were fitted to the unsmoothed data of Table 2, using weights as in (2). There is so much sampling variability in the unsmoothed data that even the linear parameters b_{ij} were often not significantly different from 0; also the error mean squares differ very little from binomial variances, so that fitting of higher-degree polynomials would be a waste of effort over periods of a month duration. Of course, the fitted models (3) cannot be extrapolated successfully outside their month; the functions are certainly not linear over the whole year. If the first order Markov chain assumption holds within a given month, then for that month we would have

$$a_{000} = a_{100},\ b_{000} = b_{100},\ a_{010} = a_{110},\ b_{010} = b_{110}.$$

The F-ratios for testing these hypotheses are given in Table 3.

Table 3. F-ratios for testing Markov chain hypotheses. Monthly regression analyses.

Month	Hypothesis	
	$a_{000} = a_{100},\ b_{000} = b_{100}$	$a_{010} = a_{110},\ b_{010} = b_{110}$
January	0.26	0.07
February	0.64	0.50
March	1.44	0.18
April	0.99	2.44
May	4.84*	0.38
June	2.17	1.28
July	2.11	0.37
August	0.06	2.43
September	0.16	6.18**
October	1.24	0.14
November	1.04	1.81
December	0.97	3.00

F-ratios shown have 2 and 52 to 58 degrees of freedom (depending on month).

* Ratio significant at 5% level.

** Ratio significant at 1% level.

We see from Table 3 and Figure 1 that only the largest observed departures from the hypotheses, in May for the comparison $\hat{p}_{000}$ with $\hat{p}_{100}$ and in September for the comparison $\hat{p}_{010}$ with $\hat{p}_{110}$, attain the conventional significance limits. But we have performed 24 F-tests, and it is doubtful whether these departures should be judged significant. The probability that 2 or more independent F-tests out of 24 exceed the 5% significance level by chance alone is 0.34.

For a somewhat different test of the Markov assumption, using contingency table techniques rather than regression analysis, we grouped the triple counts into 33 groups of 11 days each (ignoring days 1 and 2 for this exercise), thus:

$$N_{ijk}(d) = \sum_{m=0}^{10} n_{ijk}(d+m), \qquad d = 3,\ 14,\ 25,\ 36,\ \ldots,\ 365.$$

For example, the resulting totals for days 3-13 over the 107 years are

	$N_{000} = 548$	$N_{100} = 137$	$N_{010} = 92$	$N_{110} = 78$
	$N_{001} = 133$	$N_{101} = 35$	$N_{011} = 84$	$N_{111} = 70$
Total	681	172	176	148

Assuming a first-order Markov chain model with transition probabilities p_{00}, p_{01}, p_{10}, p_{11} which are essentially constant over the 11 days considered, their estimates are for these days

$$\hat{p}_{00} = \frac{N_{000}+N_{100}}{N_{000}+N_{100}+N_{001}+N_{101}} = \frac{685}{853} \div .803 \quad , \quad \hat{p}_{01} = 1 - \hat{p}_{00} \div .197 ,$$

$$\hat{p}_{10} = \frac{N_{010}+N_{110}}{N_{010}+N_{110}+N_{011}+N_{111}} = \frac{170}{324} \div .525 \quad , \quad \hat{p}_{11} = 1 - \hat{p}_{10} \div .475 .$$

The expected numbers of triples, say E_{ijk}, conditional on fixed marginal totals are then given by

$$E_{000} = 681 x \hat{p}_{00} \div 547,\ E_{100} = 172 x \hat{p}_{00} \div 138,\ E_{010} = 176 x \hat{p}_{10} \div 92,\ E_{110} = 148 x \hat{p}_{10} \div 78,$$

$$E_{001} = 681 x \hat{p}_{01} \div 134,\ E_{101} = 172 x \hat{p}_{01} \div 34,\ E_{011} = 176 x \hat{p}_{11} \div 84,\ E_{111} = 148 x \hat{p}_{11} \div 70 .$$

The goodness of fit test of the expected values E_{ijk}, calculated on the first-order Markov assumption, compared with the observed counts N_{ijk} is obtained from

$$\chi^2 = \sum (N_{ijk} - E_{ijk})^2 / E_{ijk}$$

$$\div 0.06 \text{ in the present case,}$$

the summation being over the 8 triple types and χ^2 having 2 d.f. For this group of days, the fit is obviously very good. For the complete listing of χ^2 values for all groups, see Table 4.

Table 4. χ^2 statistics for testing Markov chain hypothesis. Analysis of 11-day totals.

11 days beginning on day d	χ^2(2d.f.)	11 days beginning on day d	χ^2(2d.f.)	11 days beginning on day d	χ^2(2d.f.)
3	0.06	124	7.71 *	245	1.15
14	0.42	135	4.63	256	5.42
25	0.99	146	2.35	267	10.47 **
36	1.17	157	6.58 *	278	0.06
47	0.64	168	5.76	289	0.34
58	3.56	179	0.66	300	1.44
69	0.66	190	2.55	311	2.12
80	0.88	201	1.24	322	6.03 *
91	5.56	212	1.35	333	3.71
102	2.67	223	2.01	344	1.14
113	3.16	234	1.77	355	0.81

* Significant at 5% level, ** Significant at 1% level.

Again we must bear in mind that in 33 χ^2 tests, some may appear as significant by chance. The probability of obtaining four or more values exceeding the nominal 5% significance level, out of 33 such values, is 0.08, so that the "significant" χ^2 values above do not have their nominal significance levels. However, it must be admitted that there is some evidence that the first order model with probabilities constant over the 11 day periods is not entirely adequate statistically. Whether any deviations are important in practice is an open question, but the suggestion is that p_{100} may be below p_{000} during autumn and early winter, and p_{010} may be below p_{110} during late winter and spring, compare Fig. 1.

We next proceed to discuss whether any departures from the model seriously affect distributions of wet and dry spells.

3. ANALYSIS OF WET AND DRY SPELLS

Accepting the hypothesis that a first order Markov chain model fairly adequately fits the data, means that only two transition probabilities need be examined, namely

$$p_{00}(d) = p_{000}(d) = p_{100}(d) = \Pr(\text{day } d \text{ is dry} \mid \text{day } d\text{-1 is dry})$$

$$p_{10}(d) = p_{010}(d) = p_{110}(d) = \Pr(\text{day } d \text{ is dry} \mid \text{day } d\text{-1 is wet}).$$

The probabilities for day d being wet are, of course, the complements of these;

thus $$p_{01}(d) = 1 - p_{00}(d) \text{ and } p_{11}(d) = 1 - p_{10}(d)$$

Estimates for $p_{00}(d)$ may be obtained by a weighted average of $\hat{p}_{000}(d)$ and $\hat{p}_{100}(d)$, or more directly, by

$$\hat{p}_{00}(d) = \frac{n_{000}(d) + n_{100}(d)}{n_{000}(d) + n_{100}(d) + n_{001}(d) + n_{101}(d)} \quad \ldots(4)$$

Similarly $$\hat{p}_{10}(d) = \frac{n_{010}(d) + n_{110}(d)}{n_{010}(d) + n_{110}(d) + n_{011}(d) + n_{111}(d)} \quad \ldots(5)$$

Corresponding to the sample results in Tables 1 and 2 we have estimates as in Table 5.

Table 5. Transition probability estimates

Day of year d	$\hat{p}_{00}(d)$	$\hat{p}_{10}(d)$
1	.8472	.6176
2	.7561	.5833
3	.8182	.3000
.	.	.
.	.	.
.	.	.
189	.5111	.3387
190	.5455	.3333
.	.	.
.	.	.
.	.	.
364	.7639	.6571
365	.6923	.4828

Smoothed estimates of $p_{00}(d)$ and $p_{10}(d)$, say $\hat{\hat{p}}_{00}(d)$ and $\hat{\hat{p}}_{10}(d)$, have been computed using a 31-day weighted moving average similar to (1) and (2) above. The results are plotted in Fig. 2.

We now consider a model to fit the observed lengths of dry and wet spells, and the lengths of complete dry-wet cycles. If a first-order Markov chain with constant transition probabilities is assumed, spell lengths would have geometric probability distributions thus:

$$\Pr(\text{dry spell is of length n days}) = p_{00}^{n-1}\, p_{01}, \quad n = 1, 2, 3, \ldots \qquad \ldots(6)$$

$$\Pr(\text{wet spell is of length n days}) = p_{11}^{n-1}\, p_{10}, \quad n = 1, 2, 3, \ldots \qquad \ldots(7)$$

with means

$$\text{mean dry spell length} = \sum_{n=1}^{\infty} n p_{00}^{n-1}\, p_{01} = 1/(1 - p_{00}) \qquad \ldots(8)$$

$$\text{mean wet spell length} = \sum_{n=1}^{\infty} n p_{11}^{n-1}\, p_{10} = 1/(1 - p_{11}) = 1/p_{10}. \qquad \ldots(9)$$

Also, a dry-wet cycle consisting of a complete dry spell followed by a complete wet spell would have total length with a "generalised negative binomial" distribution:

$$\Pr(\text{cycle length} = n \text{ days}) = p_{01}p_{10}(p_{00}^{n-2} + p_{00}^{n-3}\, p_{11} + \cdots + p_{00}p_{11}^{n-3} + p_{11}^{n-2})$$
$$n = 2, 3, 4, \ldots \qquad \ldots(10)$$

These distributions were considered by Gabriel and Neumann (1957) for the winter months at Tel Aviv, and were found to give good fits. However, it must be noted that even if the first-order Markov assumption is valid but the transition probabilities are not constant over the period being considered, these distributions should not be expected to fit data very well. For instance, the probability of a dry spell which begins on day d of the year being of length 1 day is $p_{01}(d+1)$, and being of length n days is

$$p_{00}(d+1)\, p_{00}(d+2) \;\ldots\; p_{00}(d+n-1)\, p_{01}(d+n), \qquad n = 2, 3, 4, \ldots .$$

This distribution is in general neither geometric nor independent of d. This fact may explain why some authors have not obtained good fits using, for example, data for whole years. In Melbourne, we have clear evidence that the transition probabilities do vary through the year, (see Figure 2). But although 107 years is a substantial record, we do not have sufficient spells of a given type beginning on a given day to make worthwhile tests of these non-geometric distributions.

As the transition probabilities seem reasonably constant over monthly intervals, we decided to test the fit of the distributions (6), (7) and (10) on a (roughly) monthly basis. We considered those spells and cycles which began after the first day in each of 12 "months" of 30 days, approximating to calendar months, and before the 29th day of the month; these conventions were used to remove the possibility of double counting of spells across months and so that cycles (minimum length 2 days) had a chance of being completed within the month. For spells and cycles which nevertheless extended into the next month, the lengths used included such carryover except in the case of "December", days 331-360, as it was inconvenient to store in the computer more than one year's data at a time. Thus some lengths for December are biassed downwards; this bias is unfortunate because, for instance, the longest dry spell thus missed out on being fully included; it extended over 40 days in December 1954 and January 1955. Incidentally, the geometric distribution suggests that about one out of 3000 dry spells around January should be of 40 days or longer.

As a sample of the observed distributions and the expected frequencies based on (6), (7) and (10) with parameters estimated as in (7) and (8), we give in Tables 6 and 7 the results for May (more precisely, spells and cycles which began after day 121 but before day 149 in the years) and for September (after day 241 but before day 269), the months most suspected earlier of deviating from first-order Markovian assumptions.

Table 6. Spell and cycle length frequency distributions for May

Length (days)	Dry Spells Observed	Dry Spells Expected	Wet Spells Observed	Wet Spells Expected	Dry-Wet cycles Observed	Dry-Wet cycles Expected
1	247	221.21	229	222.03	-	-
2	122	139.52	136	139.73	89	77.75
3	82	87.99	85	87.94	94	97.97
4	50	55.50	62	55.34	77	92.58
5	32	35.00	31	34.83	87	77.77
6	16	22.08	18	21.92	72	61.25
7	18	13.92	11	13.79	37	46.30
8	11	8.78	10	8.68	27	34.03
9	10	5.54	3	5.46	27	24.51
10	4	3.49	6	3.44	13	17.37
11	2	2.20	1	2.16	15	12.16
12	3	1.39	3	1.36	10	8.43
13	1	.88	0	.86	3	5.79
14	1	.55	2	.54	6	3.95
15	0	.35	2	.34	1	2.68
16	0	.22	0	.21	3	1.81
17	0	.14	0	.13	5	1.22
18	0	.09	0	.08	2	0.81
19-	0	.15	0	.14	0	1.60
Total	599	599	599	599	568	568
Mean	2.71 days		2.70 days		5.41 days	

In Table 8 we show the results of calculating χ^2 goodness-of-fit statistics for the spell and cycle length distributions for each month. The statistics were calculated after grouping of certain lengths in the tails of the distributions so as not to have too low expected frequencies; this means, however, that a range of degrees of freedom have to be considered because parameters were estimated from the ungrouped data (see Kendall and Stuart (1961) p.430). As is turned out, this complication is not important. Three statistics out of 36 reached conventional significance levels showing that, in the main, the distributions (6), (7) and (10) fit the observed monthly data well. Of the exceptions, it seems possible that the April dry-wet cycle distribution fits poorly by chance, for we also calculated similar statistics for wet-dry cycles (the definition of a cycle being arbitrary as to which type of spell should come first), and the April result was quiteinsignificant (χ^2 = 13.63 with 13-15 d.f.). The August dry spell data show an excess of 4-day spells over the expected number, and the September wet spell data show (see Table 7) a deficit of 1-day spells and an excess of 2-, 3- and 4-day spells over those expected on the geometric distribution assumption. Whether these are real effects, or arise by chance amongst the many tests performed, is not known.

Table 7. Spell and cycle length frequency distributions for September

Length (days)	Dry Spells Observed	Dry Spells Expected	Wet Spells Observed	Wet Spells Expected	Dry-Wet cycles Observed	Dry-Wet cycles Expected
1	239	242.41	205	238.82	-	-
2	137	147.94	163	147.12	88	89.33
3	86	90.28	114	90.64	108	109.55
4	69	55.10	67	55.84	79	100.76
5	42	33.62	33	34.40	91	82.38
6	26	20.52	12	21.19	81	63.14
7	14	12.52	7	13.05	50	46.46
8	3	7.64	11	8.04	35	33.24
9	1	4.66	4	4.95	27	23.29
10	4	2.85	2	3.05	14	16.07
11	0	1.74	0	1.88	7	10.95
12	1	1.06	3	1.16	7	7.38
13	0	.65	1	.71	3	4.94
14	0	.39	0	.44	2	3.28
15	0	.24	0	.27	3	2.17
16	0	.15	0	.17	2	1.42
17-	0	.22	0	.26	0	2.65
Total	622	622	622	622	597	597
Mean	2.57 days		2.60 days		5.15 days	

Table 8. χ^2 goodness-of-fit statistics for spells and cycles.

Period (days)	Dry Spell χ^2	(d.f.)	Wet Spell χ^2	(d.f.)	Dry-Wet cycle χ^2	(d.f.)
January (1-30)	11.2	(16-17)	5.3	(5- 6)	17.3	(16-18)
February (31-60)	13.7	(15-16)	8.4	(5- 6)	6.7	(14-16)
March (61-90)	9.6	(13-14)	6.6	(5- 6)	21.7	(14-16)
April (91-120)	15.7	(11-12)	7.9	(8- 9)	31.4**	(13-15)
May (121-150)	13.9	(9-10)	7.6	(9-10)	18.2	(11-13)
June (151-180)	4.8	(7- 8)	9.3	(9-10)	14.0	(11-13)
July (181-210)	12.1	(7- 8)	12.4	(9-10)	12.4	(10-12)
August (211-240)	17.0*	(8- 9)	7.7	(9-10)	12.9	(11-13)
September (241-270)	14.6	(8- 9)	23.4**	(8- 9)	14.9	(10-12)
October (271-300)	5.6	(10-11)	6.3	(8- 9)	12.9	(11-13)
November (301-330)	12.0	(11-12)	8.5	(7- 8)	17.5	(11-13)
December (331-360)	9.0+	(12-13)	8.3+	(6- 7)	13.8+	(12-14)

+ Biassed data. * Significant at 5% level. ** Significant at 1% level.

It is of interest to note that we now have two methods for estimating the parameters p_{00} and p_{10}, namely from the analysis of the data using triple-type counts (it would be sufficient to use counts on pairs of consecutive days, in fact) and also from the analysis of spell lengths. For comparison purposes, we show in Table 9 the estimates using the latter method, assuming the parameters are constant over the 30 day "months" and employing (8) and (9), and also show the moving average estimates obtained by the former method, Figure 2, at the middle of the same months. Table 9 also gives the observed mean spell-, and cycle-lengths by months.

Table 9: Spell and cycle parameters.

Period (days)	Observed mean lengths (days)			Parameter estimates			
				p_{00}		p_{10}	
	Dry Spell	Wet Spell	Dry-Wet Cycles	Eqn. (8)	Fig. 2	Eqn. (9)	Fig. 2
January (1-30)	5.39	1.88	7.58	.8144	.8039	.5330	.5349
February (31-60)	4.87	1.92	7.03	.7946	.8132	.5206	.5078
March (61-90)	4.29	1.91	6.27	.7669	.7762	.5226	.5168
April (91-120)	3.48	2.52	6.05	.7122	.7181	.3966	.4065
May (121-150)	2.71	2.70	5.41	.6307	.6354	.3707	.3786
June (151-180)	2.32	2.85	5.14	.5695	.5836	.3510	.3521
July (181-210)	2.27	2.67	4.93	.5588	.5522	.3746	.3757
August (211-240)	2.45	2.73	5.23	.5926	.5829	.3658	.3625
September (241-270)	2.57	2.60	5.15	.6103	.6129	.3840	.3811
October (271-300)	2.91	2.48	5.33	.6568	.6377	.4025	.4013
November (301-330)	3.39	2.18	5.65	.7048	.6964	.4578	.4424
December (331-360)	3.77+	2.03+	5.73+	.7350+	.7385	.4932+	.4852

+ Biassed by truncation of spells at end of year

We note that although the parameters p_{00} and p_{10} have been estimated by two quite distinct methods, agreement is obtained with a $\pm .02$ variation. Thus if mean spell lengths were predicted using (8) and (9) with the Figure 2 parameters, agreement with the observed means would be very close. The slight discrepancies between mean cycle lengths and the sum of the mean dry and wet spell lengths is due to the fact that our convention for counting spells included a few more spells than complete dry-wet cycles.

4. CONCLUSIONS

(a) There is no very marked disparity between the short term patterns of wet and dry days in Melbourne and a first-order Markov chain model in which previous history only influences a given day via the immediately preceding day. The dependence of one day on the next varies throughout the year, so the model has to have non-stationary transition probabilities. The model is least successful in the months leading into and out of winter, but may still be adequate then for practical purposes. We do not claim that the process is necessarily Markovian, as other models might also fit the data adequately.

(b) Longer term patterns of dry and wet spells, and complete cycles of these, are well fitted by a geometric and a "generalised negative binomial distribution" respectively, in accord with the Markov chain model, provided the parameters of the latter remain reasonably constant for the period considered. This is so for 30-day periods, except possibly for months leading into and out of winter. We have thus allowed for seasonal effects, but we have not allowed for changes in the weather pattern over years. It would be of interest to investigate whether our transition probability estimates should vary from year to year.

(c) The above conclusions are in accord with work by Gabriel and Neumann (1957, 1962) who dealt with winter months at Tel Aviv and who had considerably less data to analyse.

Much of the data could not be published in our paper, of course. If readers require more details concerning our sample tables or other results, we will be pleased to supply them.

REFERENCES

Gabriel, K.R. and Neumann, J 1957 "On the distribution of weather cycles by length". Quart. J. R. Met. Soc. V. 83, No. 357, pp 375-380.

1962 "A Markov chain model for daily rainfall occurrence at Tel Aviv". Quart. J. R. Met. Soc. V. 88, No. 375, pp 90-95.

Kendall, M.G. and Stuart, A. 1961 The Advanced Theory of Statistics. Vol. II. Griffin, London.