

THE SPATIAL DISTRIBUTION OF MEAN AIR TEMPERATURE IN SOUTHEASTERN NEW SOUTH WALES

M. Johnson, J. D. Kalma, and J. Caprio

Division of Land Use Research, CSIRO, Canberra

(Manuscript received February 1975; revised May 1976)

ABSTRACT

This note describes the use of multiple regression techniques for estimating long-term climatic averages of mean monthly air temperature from elevation, latitude, and distance from the coast. It is shown for 22 climate stations in southeastern New South Wales that for mean maximum temperature the regression accounts for over 89 per cent of the variance in each month, and the estimates of monthly means of the diurnal temperature range account for over 85 per cent of the variance in each month.

Equations derived for southeastern New South Wales have been used to obtain detailed information on the spatial distribution of long-term averages of monthly mean maximum and minimum air temperatures for an area on the lower south coast. Good general agreement is shown to exist between estimated and observed values.

INTRODUCTION

It is generally accepted that within relatively restricted climatic zones, differences in elevation and latitude account for most of the spatial variability of monthly and annual temperature means. Other factors which may be important are landform and relief or the proximity of large water bodies.

Several North American studies have been reported in which temperature means are estimated from multiple regression equations. J.W. Hopkins (1938) examined the least squares regression of mean monthly air temperatures recorded at 44 stations in the Canadian Great Plains on latitude, longitude, and altitude. In a paper published 30 years later (J.W. Hopkins 1968) more than 200 stations were included and polynomials of moderate degree in the same three parameters were obtained. C.D. Hopkins (1960) and Lee (1969) successfully estimated mean monthly temperatures in the northeastern USA from elevation and latitude. Solomon et al. (1968) used latitude, elevation, distance from the coast, and land form as independent variables in temperature estimates in Newfoundland and Labrador.

Little Australian work has been done on developing methods for estimating temperature means, although several papers report on studies of short-term estimation of temperature extremes (e.g., Veitch 1970). Thompson (1973a) recently discussed temperature distribution in the Armidale district in relation to relief variations. Elsewhere (Thompson 1973b) synoptic classification of air flow types has been applied to mesoscale variations in temperature and rainfall in the same area.

The present paper presents an estimation of air temperature means from elevation, latitude, and distance from the coast for a network of 22 stations in southeastern New South Wales. The derived estimation equations have sub-

sequently been used to obtain detailed information on the spatial distribution of temperature in a 6000 km² region on the lower south coast of New South Wales (Fig 1), for a resource survey. Some kind of technique for temperature estimation was necessary because of the lack of climatological stations in the survey area thus providing the primary motivation for this study.

METHODS AND RESULTS

Data

Records of monthly maximum and minimum temperatures for 22 climate stations in southeastern New South Wales were obtained from the Bureau of Meteorology. A list of stations, and the period covered for each, is given in Table 1 and their positions are shown in Fig 1.

From these records, the mean monthly maximum temperature and diurnal temperature range were calculated for 12 months for each station. For each mean value the corresponding variance was also obtained.

Table 1 Data stations

Station	Period of record	Log (miles from coast +5)	Elevation (ft)	Latitude (dec.deg.)
Bega	1911-1935	1.146	44	36.67
Bodalla	1911-1935	0.845	40	36.15
Bombala	1912-1935	1.634	2450	36.90
Bondi SF	1930-1950	1.708	3000	37.15
Bowral	1911-1935	1.505	2171	34.50
Braidwood	1911-1935	1.568	2150	35.43
Burrinjuck	1912-1935	2.053	1305	35.00
Canberra ACT	1914-1935	1.887	1837	35.30
Canberra FO	1927-1935	1.887	1872	35.30
Cooma (Lambie St)	1911-1935	1.763	2664	36.23
Crookwell	1916-1935	1.914	2910	34.47
Duntroon	1912-1930	1.863	1918	35.27
Goulburn	1911-1935	1.792	2302	34.75
Jervis Bay	1911-1935	0.699	252	35.10
Kiandra	1911-1935	1.987	4578	35.88
Moruya Heads	1911-1935	0.699	35	35.92
Nalbaugh	1938-1956	1.568	2400	37.07
Nimmitabel	1911-1935	1.672	3465	36.52
Nowra (RANAS)	1957-1960	1.255	354	34.95
Nowra	1911-1935	1.146	50	34.88
Queanbeyan	1911-1935	1.839	1899	35.35
Yass	1907-1926	1.991	1626	34.85

Multiple regression equations

Using the above data, 24 weighted regression equations were obtained for the mean maximum temperature (°C) and the mean diurnal temperature range (°C) for each month of the year.

When determining each regression equation the mean monthly value for each station was weighted by a factor w , where $w = 1/(\text{variance of the mean for that station})$. This factor accounts for the temperature variation between years for an individual station. Any bias between stations, due to the slightly different periods of the records used, has been ignored. However, since only three stations

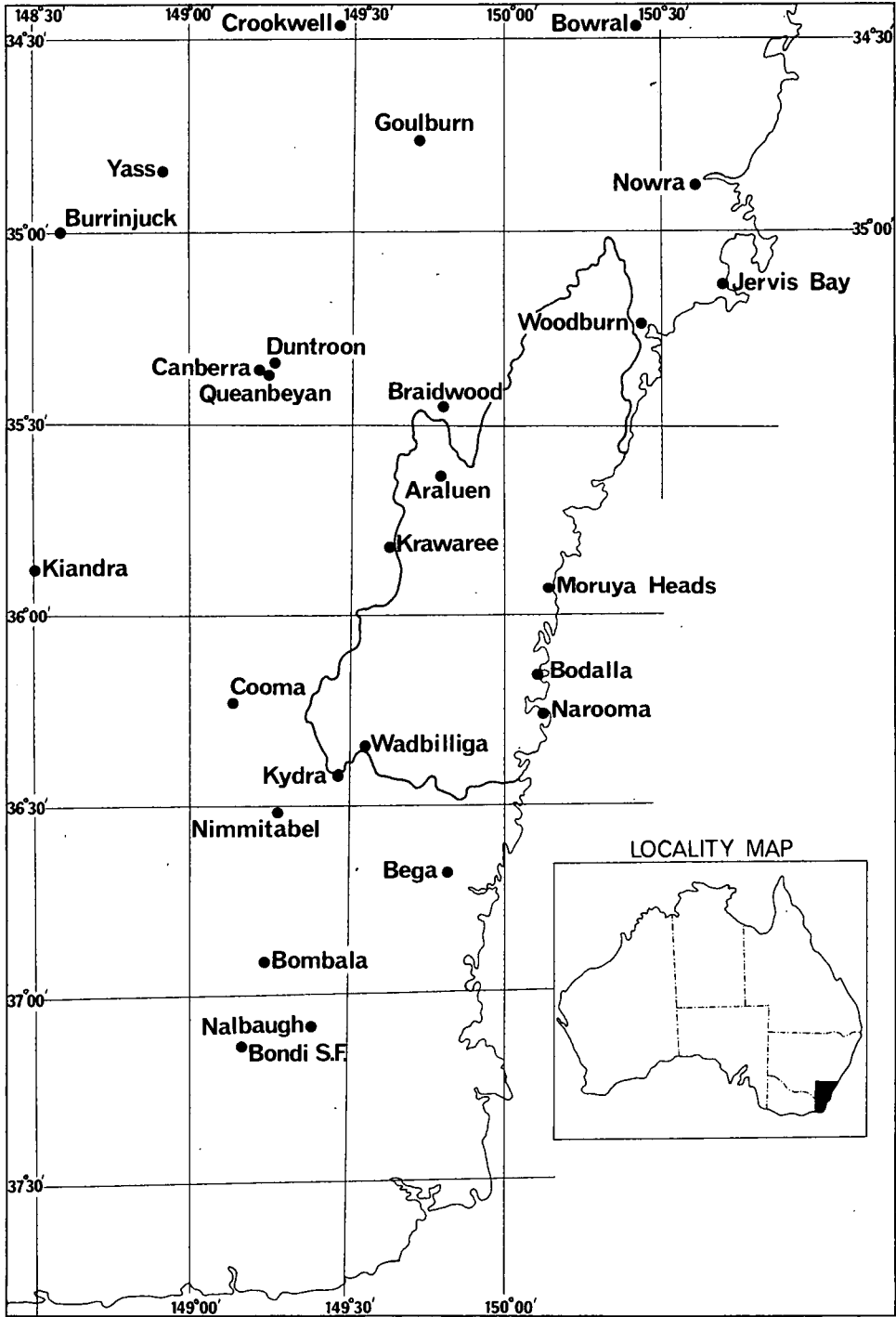


Fig 1 Southeastern New South Wales, showing the survey area, the location of the data stations, the testing stations, and other towns mentioned in the text

had records completely outside the 1911-1935 period, any influence due to long-term temperature changes should be small.

Each equation is of the form

$$Y = a + b(X1) + c(X2) + d(X3) + e(X4) + f(X5)$$

where Y = mean maximum temperature or mean diurnal temperature range
 $X1$ = $\log(\text{miles from coast} + 5) - 1.5646$
 $X2$ = elevation above mean sea level (ft) - 1787.4
 $X3$ = latitude - 35.62
 $X4$ = $(X1)^2$
 $X5$ = $(X2)(X3)$

and a-f are the appropriate regression coefficients for the month.

The numerical constants in the expressions for $X1$, $X2$, and $X3$ are the mean values of $\log(\text{miles from coast} + 5)$, elevation, and latitude, respectively, for the 22 data stations. The means were included in the expressions to reduce the correlation between the regressors. The constant (+ 5) in the expression for $X1$ was included to allow for places such as Montague Island, which is about 4.7 miles off the coast of the mainland.

Initially the six second-order terms of $X1$, $X2$, and $X3$ were incorporated in the regression equations for both temperature parameters. However, an elimination option in the computer program discarded these terms, apart from $X4$ and $X5$. This elimination is based on a t-test whereby all predictors with lowest t are successively eliminated until the smallest t-value exceeds 1.5. Because nearly all second-order terms were thus eliminated, it was decided to employ the simpler regression model given above.

Tables 2 and 3 include the final monthly values of the coefficient of determination R^2 given by the expression

$$R^2 = 1 - (\text{error mean square} / \text{total mean square}).$$

Table 2 Regression coefficients for maximum temperature ($^{\circ}\text{C}$)

Month	a	b	$-c \times 10^3$	-d	-e	$-f \times 10^3$	R^2
January	26.680	4.648	0.220	0.809	4.197	0.219	0.929
February	26.464	4.742	0.202	0.641	2.414	0.466	0.893
March	24.259	2.805	0.218	0.325	4.057	0.522	0.924
April	19.872	0.703	0.225	0.305	3.067	0.344	0.959
May	15.971	-0.127	0.239	0.091	2.530	0.225	0.984
June	13.038	-1.907	0.225	0.157	3.175	0.297	0.989
July	12.605	-2.195	0.234	0.115	3.936	0.304	0.992
August	14.382	-1.876	0.235	0.134	4.869	0.284	0.987
September	17.437	-0.925	0.218	0.241	5.002	0.461	0.932
October	20.916	-0.317	0.215	0.682	6.789	0.250	0.922
November	23.502	2.090	0.214	1.012	6.074	0.272	0.925
December	25.544	4.475	0.236	0.742	4.575	0.106	0.929

There is only a slight decrease in the accuracy of the regression equations when reducing the number of independent variables from nine to five, since for any month the maximum decrease in R^2 adjusted for a change in the number of degrees of freedom never exceeded 0.04.

Table 3 Regression coefficients for diurnal temperature range ($^{\circ}\text{C}$)

Month	a	b	-c $\times 10^3$	d	-e	-f $\times 10^3$	R ²
January	14.710	4.154	0.156	0.387	6.113	0.520	0.971
February	13.781	4.936	0.233	0.942	2.395	0.116	0.947
March	13.558	3.998	0.186	0.855	4.983	0.688	0.926
April	12.871	3.403	0.725	1.033	5.235	1.055	0.921
May	12.193	1.780	0.677	1.248	5.850	1.080	0.923
June	11.409	0.211	0.694	1.541	6.717	1.359	0.937
July	11.446	0.232	0.810	1.563	6.357	1.494	0.875
August	12.839	0.320	0.695	1.431	6.955	1.327	0.891
September	13.218	1.583	0.522	1.371	5.655	1.402	0.856
October	14.530	1.602	0.567	0.589	8.145	0.850	0.915
November	14.769	2.630	0.394	0.194	7.855	0.643	0.958
December	14.549	3.692	0.225	0.398	6.460	0.492	0.961

The 24 values of the coefficient of determination are very high, the lowest being 0.856. The values show a seasonal trend for both maximum temperature and diurnal temperature range. The values for maximum temperature are lowest in February, increase through to July, and then decrease again. For the diurnal temperature range, however, the lowest values occur in winter and the highest in summer.

Tables 2 and 3 also give truncated values of the regression coefficients a, b, c, d, e, and f for each month for the two dependent variables, indicating that the coefficients display some seasonal variation.

Seasonal variability of regression coefficients was studied in some detail by Coote and Cornish (1958) as part of an analysis of the correlation of rainfall with latitude, longitude, and altitude in South Australia. Although such a detailed analysis could also have been carried out in this case, the interpretation of the results could be misleading, because of the possible instability of the regression coefficients due to the high correlation between distance from the coast and elevation.

The problem of instability was not considered a serious one, however, since the main aim of the study was temperature estimation rather than detailed analysis of the parameters affecting temperature.

Application of the regression results

Distribution patterns of mean air temperatures were required in a 6000 km² area on the lower south coast of New South Wales, where the Division of Land Use Research, CSIRO, is currently engaged in a pilot survey of resources. This area comprises the catchments of the Clyde, Deua-Moruya, and Tuross Rivers and lies between 35^o10' and 36^o25' lat. From a climatological point of view the area is inadequately covered. Only two stations in the area (Moruya and Bodalla) could be included in the list of stations given in Table 1, which have been used for obtaining the above multiple regression equations.

A longitude-latitude grid was imposed on a map of the south coast study area. Grid intervals are 7½' on both axes. A total of 62 grid points was marked, with 24 of these being just outside the actual study area. The area of one grid cell is about 150 km². The latitude, elevation, and the distance from the coast of each grid point was determined. Then the values of the two dependent variables, mean maximum temperature and the mean diurnal temperature range, were determined for each month using the regression equations discussed above, yielding also mean minimum temperature and mean temperature.

The value of the mean maximum temperature for each month was marked at each grid point on the map and isotherms drawn. Other terrain information was only used in a minor way. This procedure was also followed for the minimum temperature and the mean, so that for each month three temperature maps were constructed.

January and July distributions of mean maximum temperature, mean minimum temperature, and mean daily temperature are shown as examples in Figs 2 and 3.

DISCUSSION

The temperature maps display certain common features as exemplified in Figs 2 and 3:

1. The isotherms run in a direction roughly parallel to the coast. This is particularly noticeable close to the coast itself.
2. The isotherms are closer together near the coast indicating a steeper temperature gradient in this region.
3. Near the extreme southwest corner of the survey area the isotherms are very curved. This feature is due to the high elevation of the area as indicated by the high elevations at Kydra and Wadbilliga. The curvature of the isotherms on the western boundary can be attributed to the variation in elevation. This explains the divergence of the isotherms near Araluen.

To evaluate the accuracy of the regression equations, estimated monthly values of maximum temperature and the diurnal temperature range were compared with average recorded values of these parameters for three stations not used in obtaining the regression equations.

The stations selected (Woodburn, Krawaree, and Narooma) give a range of values for the input variables in the regression equations, that is for latitude, elevation, and distance from the coast, as shown in Table 4.

Table 4 Testing stations

Station	Distance from coast (miles)	Elevation (ft)	Latitude (dec.deg.)
Woodburn	5.5	1200	35.40
Krawaree	31	2526	35.80
Narooma	0	100	36.23

For each station the mean maximum temperature, the diurnal temperature range, and the variance of the mean for both, were calculated for each month from temperature records. The records available, however, were incomplete and usually only five years of data were available in the period 1968-1973.

Because of the difference in the periods of the records for the data stations and the testing stations, some discrepancy would be expected between the temperature values obtained using the regression equations and the average recorded values. The magnitude of this discrepancy could not be estimated however.

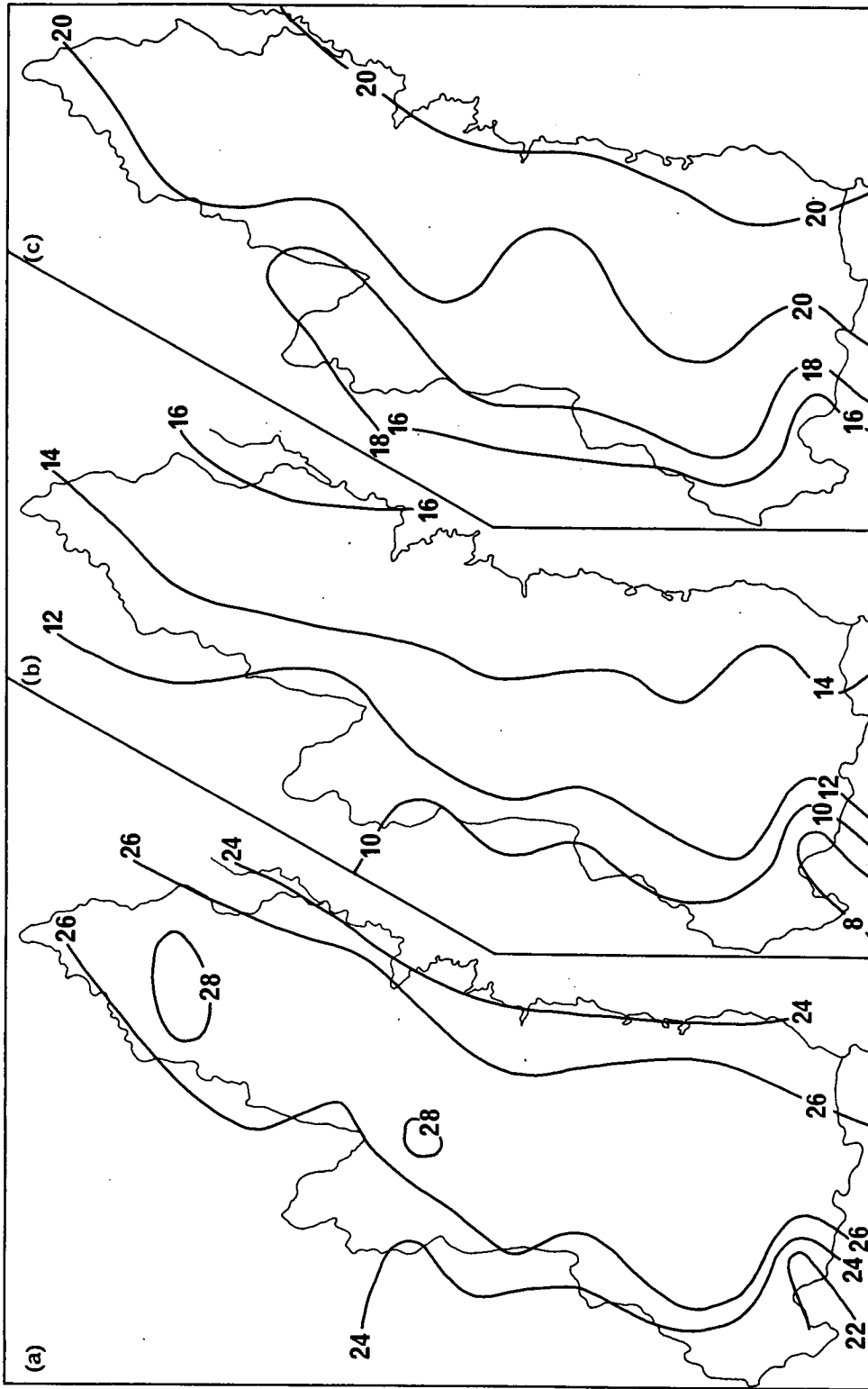


Fig 2 Distribution of (a) maximum, (b) minimum, and (c) mean monthly air temperature (°C) in January

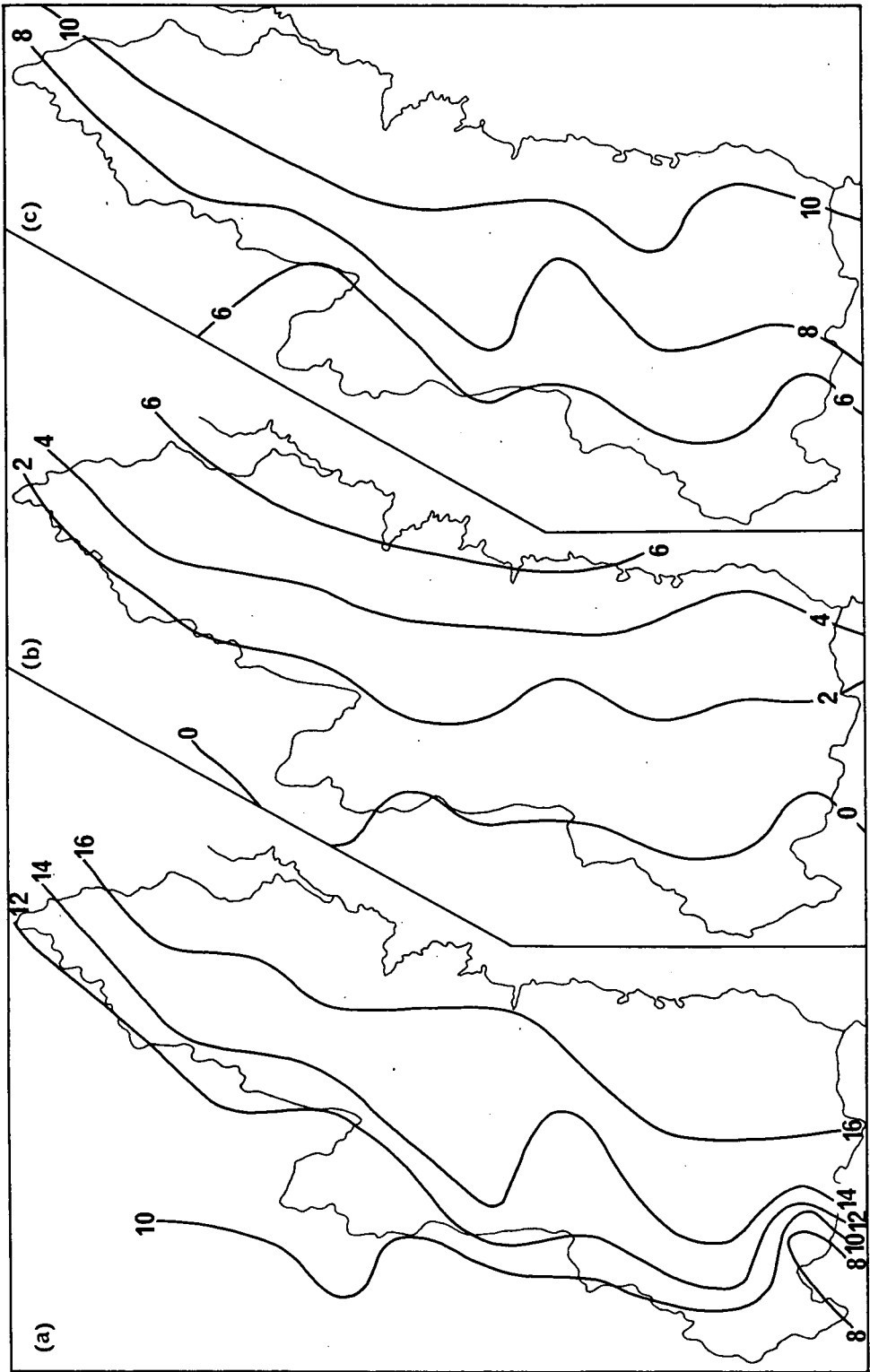


Fig 3 Distribution of (a) maximum, (b) minimum, and (c) mean monthly air temperature (°C) in July

For a particular station, month, and temperature parameter, the confidence interval (C.I.) was calculated for the estimated parameter Y_{est} at the 5 per cent confidence interval. This yields

$$C.I. = Y_{est} \pm \Delta C.I.$$

where the expression for $\Delta C.I.$ is based on work by Draper and Smith (1966; p. 122). The $\Delta C.I.$ term, however, includes the variance of the mean of the particular testing station and since this was based on only a short period of records, the resulting confidence intervals are only approximate. Furthermore the parameter ΔT was calculated, where $\Delta T = \text{mean recorded temperature} - Y_{est}$.

Based on these calculations, six graphs were drawn for the three stations and both temperature parameters. On each graph, values of $\pm \Delta C.I.$ and of ΔT were plotted for all months. These are shown in Figs 4 to 6.

The graphs show that in most cases the mean recorded temperature parameter does lie in the calculated confidence interval for that parameter and month. The agreement between the estimated and recorded temperatures varies from station to station, and also throughout the year for a particular station.

For Woodburn, for the maximum temperature, the largest discrepancies between estimated and recorded temperatures occur in the winter months, when recorded values are outside the confidence interval. In summer however the magnitude of ΔT is smaller. For the diurnal temperature range the agreement is not as good, particularly in the warmer months. In winter the estimated values are too low, and from August through to March they are too high.

For Krawaree the best agreement between estimated and recorded maximum temperatures occurs in winter. Except for July, the estimated maximum temperature is higher than the mean recorded value. The confidence intervals for the maximum temperature, and also the diurnal temperature range, are wider than those for Woodburn. The agreement for the diurnal temperature range is similar to the maximum temperature agreement. Again, except for July, ΔT is negative and the best agreement is in the winter months.

Narooma, the coastal station, gives the best general agreement of the testing stations for both temperature parameters. For both graphs the plot of ΔT lies between $\pm \Delta C.I.$ and the magnitude of ΔT is smaller than the values for Woodburn and Krawaree. The sign of ΔT does not display an obvious seasonal trend for the diurnal temperature range. As for the maximum temperature, the maximum value of ΔT occurs in November, and in general $|\Delta T|$ is somewhat higher for the diurnal temperature range.

CONCLUSIONS

Because of the variability in the results of the testing procedure outlined in the last section, it is difficult to assess the overall success of the use of a multiple regression technique in temperature estimation. From the results for the three stations, it appears that some features of the seasonal trend in temperature have not been adequately taken into account.

The results for the coastal station, Narooma, are better than those for the inland stations, Woodburn and Krawaree. This may be linked to the location of the data stations used in obtaining the regression equations, since they were either close to the coast at low altitudes, or a reasonable distance from the coast at much higher altitudes. Thus results for the inland stations may not be as good as anticipated because of their moderate distance from the coast, and moderate elevation.

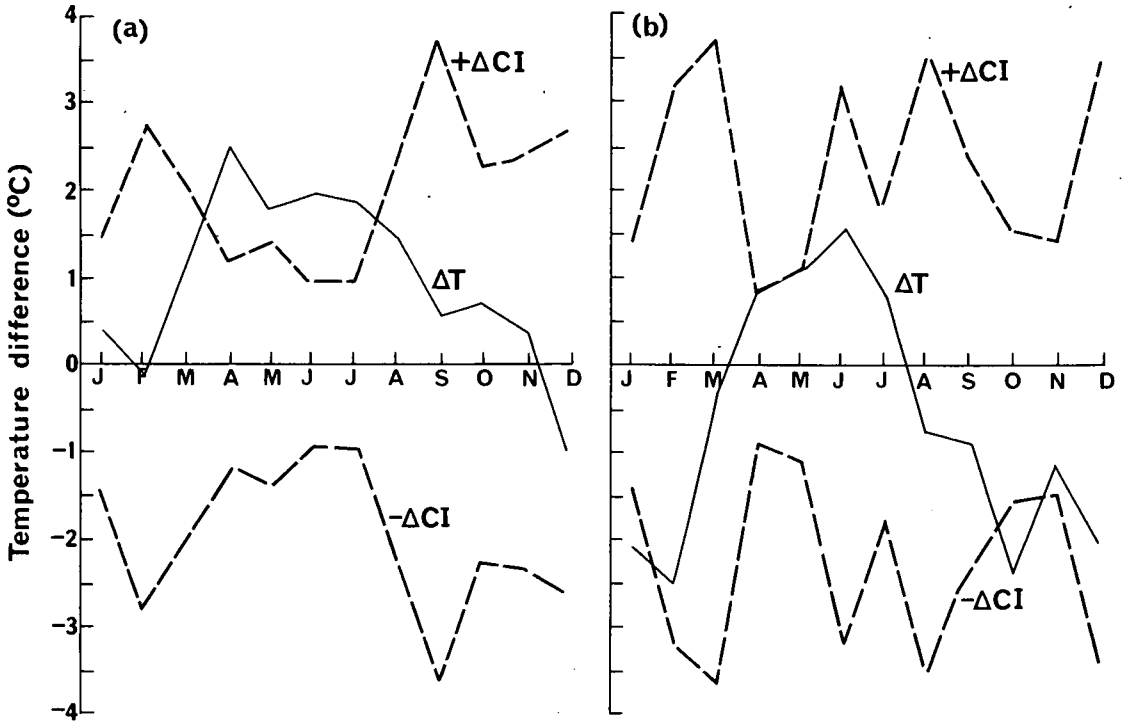


Fig 4 Graph of $\pm\Delta CI$ and ΔT against months for (a) maximum temperature and (b) for the diurnal temperature range at Woodburn

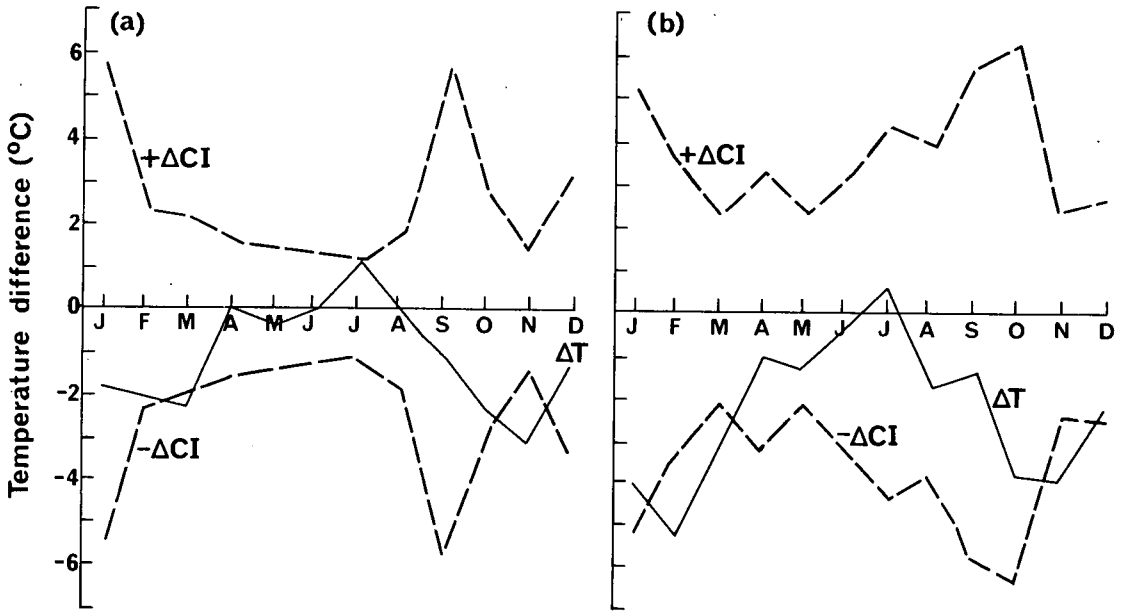


Fig 5 Graph of $\pm \Delta CI$ and ΔT against months for (a) maximum temperature and (b) the diurnal temperature range at Krawaree

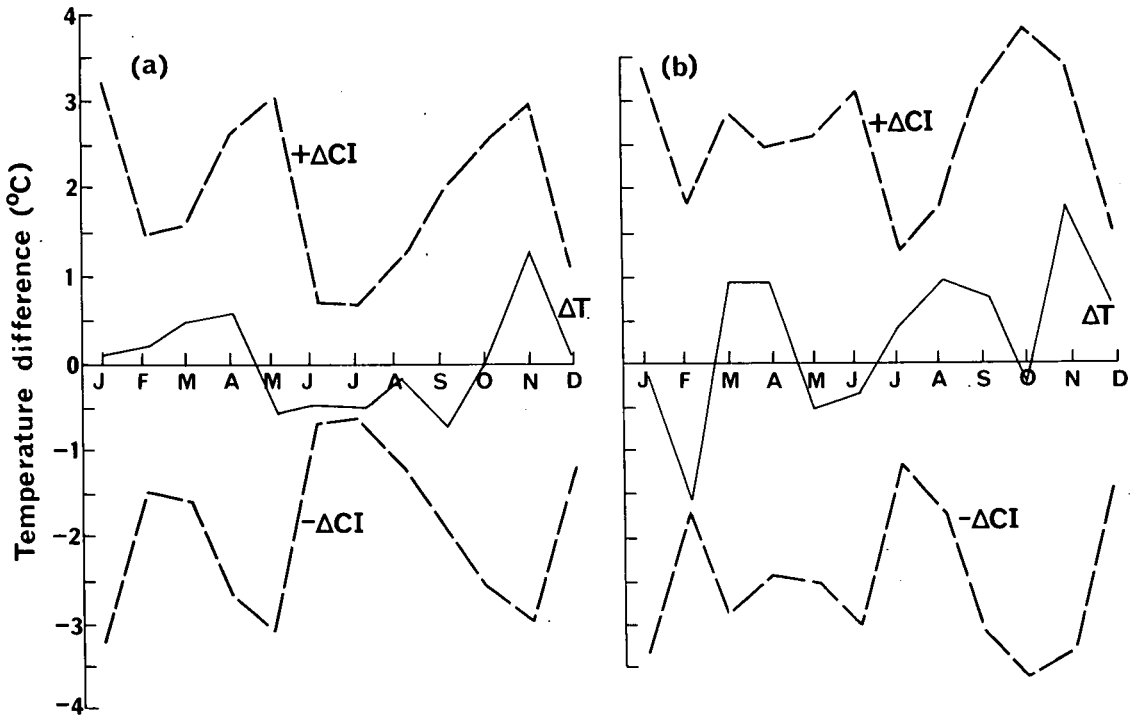


Fig 6 Graph of $\pm \Delta CI$ and ΔT against months for (a) maximum temperature and (b) the diurnal temperature range at Narooma

However, although the multiple regression technique may not always estimate the temperature for a particular month with high precision, it appears sufficiently accurate for constructing isotherms on a map of the south coast survey region. It can be used to determine the general shape of the isotherms and for assessing spatial and temporal temperature differences.

The use of a multiple regression technique also demonstrates the type of dependence on air temperature, distance from the coast, latitude, and elevation and its variation throughout the year. It is interesting to note that both maximum temperature and the diurnal temperature range display a logarithmic relationship with the distance from the coast. The elimination option described earlier indicates the importance of the product, elevation \times latitude, as well as the square of the logarithmic term. Other second order terms do not appear to be significant.

It may be concluded from the present study that multiple regression techniques appear to be a useful method for estimating maximum temperatures and diurnal temperature ranges and for demonstrating the dependence of these climatic variables on distance from the coast, latitude, and elevation. Such techniques certainly deserve continued attention and further development in regional climatic analysis when the coverage of standard climatic stations is inadequate.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Ms K. Johnston for computing and to Mr R.B. Cunningham and Mr L.G. Veitch of the Division of Mathematics and Statistics, CSIRO, for advice on statistical analysis.

REFERENCES

- Coote, G.G. and Cornish, E.A. 1958. The correlation of monthly rainfall with position and altitude of observing stations in South Australia. CSIRO Aust. Div. Math. Stat., *Tech. Pap.* 4.
- Draper, N.R. and Smith, H. 1966. *Applied Regression Analysis*. (John Wiley and Sons Inc.: New York.) 407pp.
- Hopkins, C.D. 1960. A method of estimating basin temperatures in New England and New York. *J. geophys. Res.*, 65, 3767-71.
- Hopkins, J.W. 1938. Agricultural meteorology: correlation of air temperatures in Central and Southern Alberta and Saskatchewan with latitude, longitude and altitude. *Can. J. Res.*, C16, 16-26.
- Hopkins, J.W. 1968. Correlation of air temperature normals for the Canadian Great Plains with latitude, longitude and altitude. *Can. Jnl Earth Scis*, 5, 199-210.
- Lee, R. 1969. Latitude, elevation and mean temperature in the North East. *Prof. Geogr.*, 21, 227-31.
- Solomon, S.I., Denouvilliez, J.P., Chart, E.J., Woolley, J.A., and Cadou, C. 1968. The use of a grid system for computer estimation of precipitation, temperature and run off. *Water Resources Res.*, 4, 919-29.
- Thompson, R.D. 1973a. The influence of relief on local temperatures: data from New South Wales, Australia. *Weather*, 28, 377-82.
- Thompson, R.D. 1973b. Some aspects of the synoptic mesoclimatology of the Armidale district, New South Wales, Australia. *Jnl appl. Met.*, 12, 578-88.
- Veitch, L.G. 1970. Forecasting Adelaide's maximum temperature using principal components and multiple linear regressions. *Aust. Met. Mag.*, 18, 1-21.

