Aust. Met. Mag. 34 (1986) 57-63 551.509.314

A generalised Australian Model Output Statistics system

G. A. Mills, R. G. Tapp, G. F. McNamara and D. J. Pike
Bureau of Meteorology Research Centre, Melbourne, Australia

(Manuscript received February 1986; revised May 1986)

A generalised Model Output Statistics (MOS) system for Australian applications has been developed. It is general as to location and predictand, and while a fairly comprehensive set of potential predictors is currently available, the grid-point array nature of the development data set makes it simple to add further potential predictors. The system is applied to develop seasonal equations for the prediction of maximum and minimum temperature at Australian capital cities, and it is shown that these equations are substantially more skilful than the first operational MOS equations, which were developed on an unstratified data set.

Introduction

The Model Output Statistics (MOS) technique for deriving statistical equations for the prediction of specific weather elements, using numerical weather prediction (NWP) model forecasts as input, was developed by Glahn and Lowry (1972). The method has been widely applied since that time, particularly in the United States of America, and has been articulately reviewed by Glahn (1982). Woodcock (1984) developed experimental MOS forecast equations for maximum and minimum temperature at Australian capital cities, and this work was extended to operational use for maximum and minimum temperature and probability of and quantitative precipitation forecasts for those cities by Mills and Tapp (1984) (hereafter referred to as MT) and Tapp et al. (1986a).

While Hammons et al. (1976) had demonstrated the benefits of seasonal stratification of data when generating MOS forecast temperature equations, it was found that the development data set used to derive the operational equations (MT) was of insufficient size to allow stable seasonal equations to be developed for all cities. In an attempt to reduce the effect of the annual cycle on the accuracy of the operational equations, the predictors offered were deviations from the monthly means. However, Tapp et al. (1986b), in their analysis of the errors in the first year of operational MOS forecasts, demonstrated a significant contribution to the errors of these forecasts from the failure of a single equation to account for the annual cycle of predictand variance.

Archiving of the operational prognoses has been practiced since January 1981, and combined with Woodcock's (1984) research archive from 1978-1980, inclusive, provides a data set of sufficient length that stable equations can now be expected to be obtained from seasonal stratification of these data.

A generalised set of computer programs has been developed that will enable a MOS forecast equation for any predictand to be developed for any location in Australia, provided that the predictand has been observed at that location from 1978 onwards. This paper describes the features of this system and demonstrates its use in the development of seasonal equations for the prediction of Australian capital city maximum and minimum temperature.

Design of the system

The data base and statistical technique

The essential elements of any MOS development system are a data base from which statistical relationships can be derived, and suitable statistical software to determine these relationships.

This data base has three elements. The first is the archive of Australian Region Primitive Equations (ARPE) model 24-hour prognoses (see McGregor et al. 1978), which was created in a research mode (Woodcock 1984) from 1978-80 inclusive, and has been continued operationally by the National Meteorological Centre (NMC) of the Australian Bureau of Meteorology from 1981. The second is the NMC archive of Australian region objective analyses on which the prognoses are based, and the third is the observational data which is archived by the Bureau of Meteorology. These observational data can be used both as potential predictors if valid at the time of commencement of the forecast period, or as predictands if valid at the end of the period.

There are some inhomogeneities in the prognosis archive, chiefly due to the research prognoses being prepared with a different set of physical parameterisation options to those used in the operational model. These, and their impact on the accuracy of the operational MOS forecasts, have been extensively discussed by Tapp et al. (1986b). While inhomogeneities in the climatology of a development data set will degrade the quality of statistical forecast equations derived therefrom, this effect will vary depending on the predictand in question, and, provided care is taken, the increased benefits of a longer data set may well outweigh these effects in many applications.

The statistical software used to generate the MOS forecast equations is a multiple linear screening regression package developed by Miller (1984), and has been extensively described by that author.

The design of the MOS development system

The generalised MOS development system was designed with the following criteria in mind:

(a) that a set of predictor vectors should be able to be extracted from the analysis/prognosis archive for any location in Australia simply by specifying a latitude/longitude coordinate;

(b) that the set of potential predictors should not be fixed, but that if potential relevance of another predictor can be demonstrated or postulated, then this predictor should be able to be simply added to the set;

(c) that any predictand for which an observational record exists can be merged with the predictor vectors as input to the screening regression program;

(d) that the software be sufficiently simple to operate that external users, such as Regional Offices, could generate MOS forecast equations for their own applications, with only limited training.

To accomplish these aims, a sub-grid of the full Australian region grid (see McGregor et al. 1978) was selected such that the Australian continent and Tasmania was covered by the grid, but that the large oceanic areas of the Southern and Indian Oceans were not retained. This sub-grid consists of 21 points (east-west) by 15 points (north-south) and is thus approximately one-third of the full 24 by 39 array. This means that the entire analysis and prognosis archive from 1978 to the time of writing this paper can be stored on a single magnetic tape, while retaining the grid-point array information so that a predictor vector can be generated for any location on that grid by interpolation. Thus, a predictor data set can be generated by mounting only one tape, while the grid-point array nature of the archived data will enable a user to derive any predictors which can be calculated from the grid values. Implicit is the assumption that predictors well away from the location of the predictand will not be relevant; for MOS forecasts valid near the prognosis validity time this should not be a significant restriction.

The potential predictors selected or calculated for the application in this paper are listed in Table 1. They differ from those used by MT in that only the '9-point' smoothed values were used (i.e. an average value of the potential predictor over a 3 x 3 array with its mid-point at the station location), as it has been shown by Tapp (1983) that these were more skilful predictors than were the 1 or 5-point values. An exception to this was made for the 1000 mb values of the geopotential height, temperature and dew-point, where both 1- point and 9-point values were offered, as 1-point values were occasionally selected in minimum temperature forecast equations. Several other parameters which potentially were physically relevant to a wide range of possible meteorological predictions were added to the predictors previously offered. These were: dew-points; temperature and moisture advection; vorticity and divergence; layer geopotential thickness; relative humidity lapse; mean lower tropospheric relative humidity; the total-totals stability index; the west-east and north-south gradients of mean sea level pressure; and forecast 24-hour pressure change at the station location and at two grid-points (approximately 500 km) north, south, east, and west of the station. All these variables were converted to deviations from their monthly means, and in addition the sine and cosine of the day of the year were offered. No observational data were included, as it was found that under operational conditions the number of forecasts issued was being adversely affected by the occasional failure to receive the surface observation messages in real time.

Some of these variables were selected on the basis of limited experiments assessing causes of error in the operational MOS forecasts, and others from physical intuition. The variables used here are all valid at the location for which the forecasts were to be generated. It is important to note, though, that the grid-point-array form of the development data set enables any other desired predictors to be simply calculated and added to the predictor vector. This may include data from several grid-points distant from the forecast location if such information was thought to be potentially relevant to the prediction of the element being considered.

Application to maximum and minimum temperature forecasting

The system was applied to generate MOS forecast equations for maximum and minimum temperatures for Adelaide, Brisbane, Canberra, Hobart, Melbourne, Perth and Sydney. The dependent data set consisted of the six years of daily 24-hour prognoses and the 2300 GMT analyses on which the prognoses were based, from 1978-1983 inclusive. The predictands were the maximum temperature which occurred during the day following prognosis base time, while the minimum temperature was the overnight minimum preceding the daytime maximum. Thus, the observed maxima and minima were approximately 6 hours after and 3 hours before prognosis validity time, respectively. Separate equations were generated for summer (December-February), autumn (March-May), winter (June-August) and spring (September-November). In addition, each equation was generated twice, once with the mid-season months of 1979 withheld as

Table 1. List of potential predictors offered to the screening regression package. Asterisks indicate those meteorological parameters which were calculated from the archived elements.

Geopotential height	levels	1000 mb 850 mb 700 mb 500 mb	from ANALYSIS and PROGNOSIS
Temperature			
Dew-point			
West-east wind			
South-north wind			
Relative humidity*			
Moisture advection*	levels	850 mb 700 mb 500 mb	
Temperature advection*			
Vorticity*			
Divergence*			
Geopotential thickness*	layers	1000-850 mb 850-700 mb 700-500 mb 1000-500 mb	
Relative humidity lapse*			
Mean sea level pressure			
East-west MSLP gradient*			
South-north MSLP gradient*			
Total-totals index*			
Vertically averaged relative humidity*			
1000 mb geopotential	'1-point' value		
1000 mb temperature			
1000 mb dew-point			
1000 mb constant lapse temperature*			
Precipitable water			from ANALYSIS
Saturation deficiency			
Precipitation			from PROGNOSIS
Isallobars*			
Sine of day of year			
Cosine of day of year			

independent data for verification purposes, and once with the mid-season months of 1981 withheld. This doubling of the equations developed has several benefits — it can aid in assessing the stability of the equations being generated, and also reduces somewhat the effects of year to year synoptic variability on the verification statistics of the independent data forecasts. In addition, this selection of years gave one year of independent data from the research part of the prognosis archive, and one year from the operational part of the archive.

The equations

As is frequently discussed (e.g. Glahn 1982) the selection of an appropriate 'stopping point' in multiple linear regression equation development is not a clear-cut decision. That is, when does one stop selecting further predictors to add to the equation. Three different criteria were merged to select the appropriate number of predictors for each equation in this study. First, the point beyond which the addition of a further predictor did not increase the explained variance in the dependent data by more than one per cent was noted. Second, the sets of predictors from the two equations with differing independent data were compared, and the point at which the selected predictors diverged was noted. Finally, explained variance versus number of predictors for the two months of independent data was examined, and a point of diminishing returns was chosen. These criteria were then compared, and a decision made on how many predictors to use in each equation. Frequently, this decision was clear-cut, however on occasion some subjectivity was needed, and in these cases a conservative decision was made; that is, less rather than more predictors were chosen, with some emphasis placed on their physical relevance. While this procedure may not be statistically rigorous, Glahn (1982) has stated ' . . . almost any practical stopping procedure is quite adequate . . . ', and it is felt that this procedure provides a sound compromise between statistical rigor and scientific judgement.

Fifty-six equations (seven cities, four seasons, two predictands) were generated, with the number of predictors ranging from two to seven, the median number being four. It was encouraging to note that of the 113 predictors selected for the 28 maximum temperature equations, 39 were 'new' predictors; that is, predictors that were not offered during development of the operational equations. The corresponding figures for the 28 minimum temperature equations were 110 and 45.

Table 2 lists the nine most frequently selected predictors for the maximum temperature equations and the eight most frequently selected predictors for the minimum temperature equations, with their frequency of selection. For the maximum temperature equations, the most frequently chosen predictors are a forecast low-level temperature, a forecast low-level wind component (or its geostrophic wind equivalent) and a forecast lower tropospheric relative humidity. The predictors most frequently chosen for the minimum temperature equations are an analysed (i.e. persistence) near-surface temperature, a forecast near-surface temperature, and a forecast lower tropospheric dew-point.

Independent verification

As was mentioned previously, two separate sets of equations were generated for each predictand, with the mid-season months of 1979 and 1981 respectively withheld for independent verification. These separate verifications have been combined for this discussion, and will be compared with the equivalent statistics computed by MT.

Table 3 shows the root mean square (RMS) error for maximum and minimum temperature forecasts for the mid-season months of 1979 and 1981 combined for each city and season for both the new seasonal equations and the old operational (all-year) equations, together with the differences between the two. In almost all cases there is an increase in skill in going to these new equations, with the increase tending to be greatest in those cities where the performance of the operational equations was worst relative to the forecasts of the Regional Forecasting Centres (see MT). In addition, the gains are greatest in the 'extreme' seasons of summer and winter, particularly in the southern mainland cities. It is also worth noting that there is greater impact from the new equations on the forecasts of maximum temperature than there is on the forecasts of minimum temperature, possibly due to the significant persistence component in the minimum temperature equations.

RMS error statistics for forecasts issued in the same eight months by the Regional Forecasting Centres (RFCs) at 1700 h Local Standard Time (LST) are included in Table 3 for comparison with the MOS forecasts, which are typically available by approximately 1300 LST (1100 LST in Perth). Treated as independent forecasts the MOS predictions for these months showed a general level of skill greater than or equivalent to that of the forecasts issued operationally by the RFCs, other than for Adelaide maxima.

Table 2. List of the most frequently selected predictors chosen for the maximum and minimum temperature equations, together with their frequency of occurrence (C.L. refers to constant lapse extrapolation — see MT).

Predictor	*Maximum temperatures*	*Frequency of occurrence*
Prognosis	1000 mb temperature (C.L.)	14
Prognosis	1000 mb S-N wind	9
Prognosis	S-N MSL pressure gradient	8
Prognosis	1000-500 mb averaged relative humidity	7
Prognosis	E-W MSL pressure gradient	7
Prognosis	1000-850 mb geopotential thickness	6
Prognosis	700 mb S-N wind	6
Prognosis	500 mb geopotential height	5
Analysis	1000-500 mb geopotential thickness	5
	Minimum temperatures	
Prognosis	850 mb dew-point	13
Analysis	1000 mb temperature (1 pt)	10
Analysis	1000-850 mb geopotential thickness	8
Prognosis	1000 mb temperature (9 pt)	7
Analysis	1000 mb temperature (9 pt)	5
Prognosis	1000 mb temperature (1 pt)	4
Prognosis	1000 mb temperature (C.L.)	4
Prognosis	1000-850 mb thickness	4

Table 3. Root mean square (RMS) temperature errors (C) for maximum and minimum temperature forecasts for the mid-season months of 1979 and 1981 combined for MOS forecasts using seasonal (NEW) and operational (all year) equations (OLD), together with the differences between the two, and figures for forecasts issued by the Regional Forecasting Centres (RFC) during the same months.

		Maximum temperatures				*Minimum temperatures*			
		RFC	NEW	OLD	OLD-NEW	RFC	NEW	OLD	OLD-NEW
Adelaide	Jan	4.06	3.64	4.07	0.43	2.29	2.23	2.40	0.17
	Apr	1.82	2.75	2.85	0.10	2.11	2.05	2.13	0.08
	Jul	1.17	1.47	2.03	0.56	2.08	1.74	1.76	0.02
	Oct	2.13	2.90	2.76	−0.14	2.18	2.02	2.05	0.08
Brisbane	Jan	2.03	1.36	1.51	0.15	1.08	0.78	0.88	0.10
	Apr	1.65	1.35	1.13	−0.22	1.16	0.93	1.06	0.13
	Jul	1.87	1.51	1.86	0.35	1.53	1.43	1.43	0
	Oct	2.83	2.03	1.96	−0.07	1.80	1.46	1.57	0.11
Canberra	Jan	2.80	3.06	3.53	0.47	2.22	1.73	2.02	0.29
	Apr	1.90	1.59	1.60	0.01	2.75	2.17	2.46	0.19
	Jul	2.02	1.69	2.06	0.37	2.38	2.18	2.76	0.58
	Oct	2.98	2.34	2.57	0.23	2.58	2.52	2.78	0.26
Hobart	Jan	3.11	2.83	2.97	0.14	2.35	1.76	1.85	0.09
	Apr	2.30	2.09	2.14	0.05	1.80	2.09	2.01	−0.08
	Jul	1.86	1.28	1.29	0.01	2.03	1.94	1.91	0.03
	Oct	2.67	2.24	2.28	0.04	2.20	1.99	2.11	0.12
Melbourne	Jan	3.98	3.97	5.04	1.07	2.14	1.59	1.91	0.32
	Apr	2.34	2.40	2.63	0.23	1.54	1.75	1.87	0.12
	Jul	1.76	1.46	2.27	0.81	1.61	1.52	1.84	0.32
	Oct	2.26	2.63	2.78	0.15	2.06	1.64	1.72	0.08
Perth	Jan	3.49	3.43	4.21	0.78	2.08	1.75	2.11	0.36
	Apr	2.16	2.43	2.57	0.14	1.51	1.51	1.66	0.15
	Jul	1.55	1.31	2.14	0.83	1.63	1.51	1.62	0.11
	Oct	2.66	2.82	3.14	0.32	2.05	2.16	2.31	0.15
Sydney	Jan	2.25	2.09	2.27	0.18	1.11	1.20	1.22	0.02
	Apr	2.00	1.80	1.73	−0.07	1.64	1.40	1.28	−0.12
	Jul	1.68	1.26	1.54	0.28	1.63	1.05	1.23	0.18
	Oct	2.46	1.69	2.02	0.33	2.10	1.51	1.73	0.22

Table 4. Number of large (5 °C or greater) forecast errors for maximum temperature for the eight months of independent data for seasonal and operational MOS equations, together with a seasonal stratification of these data for the seven cities combined. Figures for forecasts issued by Regional Forecasting Centres (RFC) during the same months are included for comparison.

	Seasonal	*Operational*	*RFC*
Adelaide	18	20	21
Brisbane	1	3	14
Canberra	12	14	11
Hobart	4	7	18
Melbourne	20	33	18
Perth	18	31	21
Sydney	3	4	10
Summer	41	68	56
Autumn	12	14	12
Winter	1	9	6
Spring	22	21	39

Table 4 shows the number of large (5 °C or greater) forecast errors for the eight months of independent forecasts of temperatures generated using the seasonal and operational equations for each city, together with a seasonal breakdown of these errors for all seven cities combined. A considerable reduction has been achieved, with this being greatest in Melbourne and Perth. The skill of the MOS forecasts in this regard is at least comparable with, and in many cases greatly superior to, that of the subjective forecasts issued operationally for all cities and for each season.

It was demonstrated by Tapp et al. (1986b) that a significant contribution to the errors of the operational MOS temperature forecasts in their first 12 months was due to the failure of the all-year equations to adequately reflect the seasonal variation of the variance of the maximum and minimum temperatures, particularly in the southern cities where this effect is largest. Table 5 shows the forecast standard deviations (SD) for the seasonal and

Table 5. Standard deviation (SD) of maximum and minimum temperature forecasts for seasonal (NEW) and operational (OLD) MOS forecast equations, together with the observed (OBS) SD. Data are for the mid-season months of 1979 and 1981.

		Maxima equations			*Minima equations*		
		NEW	OLD	OBS	NEW	OLD	OBS
Adelaide	Jan	3.85	2.90	5.28	3.10	2.90	3.81
	Apr	2.83	2.70	3.43	2.27	2.00	2.91
	Jul	1.77	2.85	2.44	1.66	1.27	2.95
	Oct	4.25	4.07	4.80	2.18	2.48	2.85
Brisbane	Jan	1.33	1.51	1.69	0.96	1.18	1.27
	Apr	1.79	1.66	1.93	1.59	1.46	1.88
	Jul	1.30	1.68	2.07	1.82	1.56	1.53
	Oct	1.76	1.47	2.95	1.62	1.62	2.11
Canberra	Jan	2.64	2.31	4.15	1.61	1.75	2.11
	Apr	2.28	2.47	2.36	2.89	2.79	3.88
	Jul	1.25	1.69	1.95	1.64	1.88	2.95
	Oct	3.23	3.26	4.24	2.38	2.42	3.06
Hobart	Jan	2.56	2.24	3.78	1.41	1.69	2.24
	Apr	2.70	2.59	3.36	1.97	1.81	2.55
	Jul	1.79	2.05	2.43	1.82	1.53	3.01
	Oct	2.60	2.32	3.57	2.14	1.87	2.48
Melbourne	Jan	4.43	2.77	6.64	2.38	2.22	2.92
	Apr	2.98	2.69	3.52	2.19	1.77	2.67
	Jul	1.21	2.14	1.60	1.43	0.96	2.05
	Oct	3.42	3.12	4.55	2.12	1.99	2.62
Perth	Jan	3.65	2.87	5.22	2.69	2.60	3.22
	Apr	2.12	2.43	3.12	1.88	1.57	2.25
	Jul	1.22	2.28	1.86	1.79	1.63	2.48
	Oct	3.55	3.35	4.46	2.11	2.02	3.04
Sydney	Jan	2.34	2.14	2.68	1.22	1.35	1.52
	Apr	1.65	1.80	2.62	1.61	1.68	2.32
	Jul	1.62	1.79	1.66	1.23	1.11	1.58
	Oct	2.79	2.09	3.40	1.81	1.87	2.50

operational forecasts, together with the observed SD for the mid-season months of 1979 and 1981. The seasonal equation forecasts better represent the seasonal variation of SD of the maximum temperatures at Adelaide, Hobart, Melbourne, Perth and Sydney, in particular, and of minimum temperatures at Adelaide, Hobart, Melbourne and Perth. These changes largely coincide with the improvements in RMS errors shown in Table 3.

Conclusions

The experimental MOS forecast system developed for Australian cities by Woodcock (1984) and adapted for operational use by Mills and Tapp (1984) has been generalised with a larger development data set, and the flexibility to vary the potential predictors, the potential predictand, and the location in Australia for which a forecast equation is required. A longer development data set than was available to the earlier studies is now available, enabling stable MOS equations to be obtained from seasonal stratifications of these data.

It has been demonstrated that seasonal stratification of the development data results in equations for the prediction of maximum and minimum temperatures at Australian capital cities which produce, in general, lower RMS errors and improved annual cycles of forecast variance than do the original MOS forecast equations. These new equations could be simply implemented operationally to provide improved guidance to Regional Forecasting Centres. Prediction equations for any other observing station could be generated as easily as they can for the Regional Forecasting Centres.

References

Glahn, H. R. 1982. Statistical weather forecasting. In *Probability, statistics and decision making in the atmospheric sciences.* A. H. Murphy and R. W. Katz (eds), Westview Press, Boulder, Colorado.

Glahn, H. R. and Lowry, D. A. 1972. The use of model output statistics (MOS) in objective weather forecasting. *Jnl appl. Met., 11,* 1202-11.

Hammons, G. A., Dallavalle, J. P. and Klein, W. H. 1976. Automated temperature guidance based on three-month seasons. *Mon. Weath. Rev., 104,* 1557-64.

McGregor, J. L., Leslie, L. M. and Gauntlett, D. J. 1978. The ANMRC limited area model: consolidated formulation and operational results. *Mon. Weath. Rev., 106,* 427-38.

Miller, A. J. 1984. Selection of subsets of regression variables. *J. Roy. Statist. Soc., A147,* 389-425.

Mills, G. A. and Tapp, R. G. 1984. The Australian operational MOS forecast system. *Tech. Report 56,* Bur. Met., Australia, 50 pp.

Tapp, R. G. 1983. A statistical comparison of concurrent analysis, prognosis and observational data for seven Australian cities. *Tech. Report 52,* Bur. Met., Australia, 30 pp.

Tapp, R. G. , Woodcock, F. and Mills, G. A. 1986a. The application of model output statistics to precipitation prediction in Australia. *Mon. Weath. Rev., 114,* 51-61.

Tapp, R. G., Mills, G. A. and McNamara, G. F. 1986b. A diagnostic study of the first year of operational MOS forecasts in Australia. *Aust. Met. Mag., 34,* 43-56.

Woodcock, F. 1984. Australian experimental model output statistics forecasts of daily maximum and minimum temperature. *Mon. Weath. Rev., 112,* 2112-21.