Climate Dynamics (2001) 17: 947–963

B. Timbal · B. J. McAvaney

An analogue-based method to downscale surface air temperature: application for Australia

Received: 14 December 1999 / Accepted: 10 January 2001

Abstract A statistical model (SM) has been developed to downscale large-scale predictors given by general circulation models (GCMs); emphasis has been put on local surface air temperature in two areas of interest: the south-west corner (SWC) of Western Australia and the Murray-Darling Basin (MDB) in southeastern Australia. This is a complementary approach to the dynamical modelling of climate change using high-resolution nested regional models. The analogue technique was chosen for this study as it has proven successful in the past for mid-latitude climate and, in particular, for forecasting in Australia. Furthermore, the analogue technique is successful in reproducing spells of anomalous events. The development and validation datasets used for both predictors and predictants cover the 1970–1993 period. Predictors are extracted from a dataset of operational analyses for the Australian region. Several predictors have been assessed alone and combined. Mean sea level pressure and temperature at 850 hPa have been identified as the most useful combination. Predictants have come from quality controlled stations with daily temperature extremes for the 1970–1993 period. Twenty two stations in the SWC and 29 in the MDB have been selected. The sensitivity of the SM has been tested to several internal parameters. The number of atmospheric predictors and the geographical domain on which large-scale fields are used are key factors that maximise the skill of the SM. Several metrics have been tested taking into account the state of the predictors on the day or, in order to describe the evolution of the atmosphere, over several days. This latter has been particularly useful in improving the representation of anomalous spells as it partially incorporates the auto-correlation of surface temperature. The correlation obtained between the observed local temperature series and the reconstructed series ranges between 0.5 and 0.8. Best results are obtained in summer and for maximum temperature. The reproduction of spells is satisfactory for most stations. The SM is then applied to large-scale fields obtained from a GCM forced by observed sea surface temperature; the improvement gained when using the SM instead of relying on the surface temperature calculated by the GCM is shown.

B. Timbal (✉) · B. J. McAvaney
Bureau of Meteorology Research Centre,
Melbourne, Australia
E-mail: b.timbal@bom.gov.au

1 Introduction

Since anthropogenic climate change has become an issue, most research has been directed toward understanding the impact on the entire climate system. Due to the complexity of the task, state-of-the-art models are most accurate for large-scale features, but are not yet suitable to describe local effects. However, most impacts are related to regional scales. Current coupled atmosphere-ocean general circulation models (CAOGCMs) provide transient scenarios of climate evolution following expected emission scenarios provided by the Intergovernmental Panel for Climate Change (IPCC) (Houghton et al. 1996). While these are broadly consistent in their predictions on the large scale, due to limited horizontal resolutions, amongst other factors, this approach has so far failed to provide consistent estimates of climate change on regional scales.

Several methods have been employed to increase atmospheric model resolution. Time-slice experiments, where high-resolution atmospheric models are forced by coarse oceanic anomalies derived from CAOGCMs are often used (Stephenson and Held 1993; Timbal et al. 1995a). Numerous integrations can be performed to investigate regional uncertainties due to model formulation or oceanic forcings (Timbal et al. 1995b). Up to now, the most common dynamical approach has been to run a regional climate model at high resolution nested within a GCM at coarse resolution. This approach has had numerous applications (Giorgi et al. 1990; a review

by Giorgi and Mearns 1999) but is not without its drawbacks (e.g. the one-way nesting and spurious boundary effects), but so far it has allowed the finest resolution for a dynamical approach.

While observations used for environmental impact modelling are usually point specific, model variables are an integration of a broad area represented by a model grid box (von Storch 1995). Statistical approaches are an alternative for bridging the scale difference (''downscaling'') between coarse grid CAOGCMs and the fine temporal and spatial scales required for regional and environmental impact studies of climate change. Statistical downscaling techniques originated in weather prediction but are commonly used in climate studies. They give access to finer resolutions than deterministic modelling, matching the scales needed for climatic change impact studies.

Several types of functions, called transfer functions, linking large-scale patterns to local surface predictant values have been used (Wilby and Wigley 1997). Most are based on large scale atmospheric predictors (e.g. mean sea level pressure, geopotential height or atmospheric thickness) for which GCMs are considered to be ''credible'' or at least more reliable than surface observations averaged over grid boxes. Predictor series are then obtained from these functions with a high temporal distribution (e.g. daily) and are more trustworthy than from a dynamical approach. Hence, climatic issues such as extreme events can be investigated (e.g. frost days, anomalous spells and return of extremes). This is a key area of research where insufficient effort has been devoted so far, as noted by the second IPCC scientific assessment (Houghton et al. 1996).

Whatever transfer function is developed, a primary step is to ensure that the statistical model has skill. In particular, the reconstructed temporal series given by the statistical approach must be more realistic then raw variables given by the GCM (Palutikof et al. 1997).

For this approach to succeed, the GCM must represent the large-scale climate realistically. If not, the lack of reasonable estimates of the predictors dooms the statistical approach to failure. At this point it is worth emphasising that, for the reasons stated, statistical approaches will always be complementary to dynamical modelling.

Australia is a particularly suitable region for the application of downscaling methods, since it has been demonstrated that climate variability and global changes have a strong impact on regional ecological and economic fluctuations (Nicholls 1997). Therefore two areas of interest have been selected according to their importance in term of agricultural production. One is the Murray-Darling Basin (MDB) and the other one is the south west corner (SWC) of Western Australia (Fig. 1). The latter is of particular concern as it appears to be affected by transient changes in long term climatologies (Allan and Haylock 1993).

In this work, firstly the datasets used to develop and validate the statistical model (both predictors and predictants) are described, as well as the GCM simulation on which the transfer function is applied. The development of the model is described, in particular the methodology and the treatment of predictors. The model is then validated using atmospheric analyses and finally applied using GCM predictors. The final section assesses the results against the aims and discusses future perspectives for the downscaling system.

2 Data

To build and validate the statistical model (SM), a long set of data is mandatory for both predictors and predictants. The database for predictors, METANAL, are historical synoptic analyses obtained from the Bureau of Meteorology operational numerical analysis system. From 1970 to 1993, at 00*UTC* and 12*UTC*, grid point analyses for the Australian region have been archived (Seaman et al. 1977). Analyses are not available beyond 1993. The series has been projected on a coherent and regular grid over the entire period. With a resolution of 1.5° on a domain from 50°S to 10°S and from 90°E to 170°E, the analyses provide a detailed description of the synoptic situation over Australia and the major air masses characterising the weather.

The analyses available contain mean sea level pressure (MSLP), geopotential heights at various pressure levels (1000, 850, 700, 500 and 250 hPa), temperature at 850 hPa (T_{850}) and wind speed at 500 hPa (Wind_{500}). Although a reasonable representation of the dynamics of the atmosphere can be inferred from these variables, no information is available on the hydrological cycle. Therefore, the moisture availability in the atmosphere affecting cloudiness and daily temperature range (Power et al. 1998) is only indirectly inferred. Consequently, more variables were extracted from the original dataset only for the subset of the time interval (namely 1978 to 1993) for which they could be retrieved. The extra variables extracted were the dew point temperature at 850 hPa (from which the relative humidity (*RH*) at the same level can be retrieved) and the vertically integrated precipitable water (*PW*). The latter is considered more reliable than the former, as moisture is integrated over the atmospheric column.

The bench mark simulation used to test the SM is an AMIP-type integration (1979–1988) of the AGCM developed by the Bureau of Meteorology Research Centre (BMRC). The AGCM is a state-of-the-art atmospheric model which has been thoroughly described in the literature. For details on the model genesis see McAvaney et al. (1978). The model configuration is described by McAvaney and Colman (1993). Since 1993, the major changes are the introduction of the Tiedtke (1989) convection scheme (McAvaney et al. 1995); improvements in surface flux formulation (McAvaney and Hess 1996) and horizontal and vertical diffusion. The resolution is R31L17. From the 17 vertical levels of the model, variables interpolated to 1000, 850 and 500 hPa were used. The horizontal resolution of the model over Australia is approximately 250 km in latitude by 350 km in longitude giving 25 and 40 point respectively in the MDB and SWC. This is not a climate change experiment, and thus at this stage we are not downscaling a climate change scenario. However, forcing the model with observed SST limits climate drift, and it is anticipated that local atmospheric variables are more realistic with annually varying SSTs. Therefore, the ability to reconstruct local temperature series using GCM output, validated against the coarse grid surface series given by the GCM, is a mandatory step in the development of the SM.

The predictants are 2 m surface-air temperature daily extremes (T_{max} and T_{min}). Most of the 29 and 22 stations (Fig. 1) used in the MDB and the SWC are part of a set of high-quality synoptic stations. The methodology used to correct and homogenise these datasets at a daily time scale has been described in Trewin and Trevitt (1996). The two regions chosen are of major interest for climate research; this is also where most of Australia's agricultural

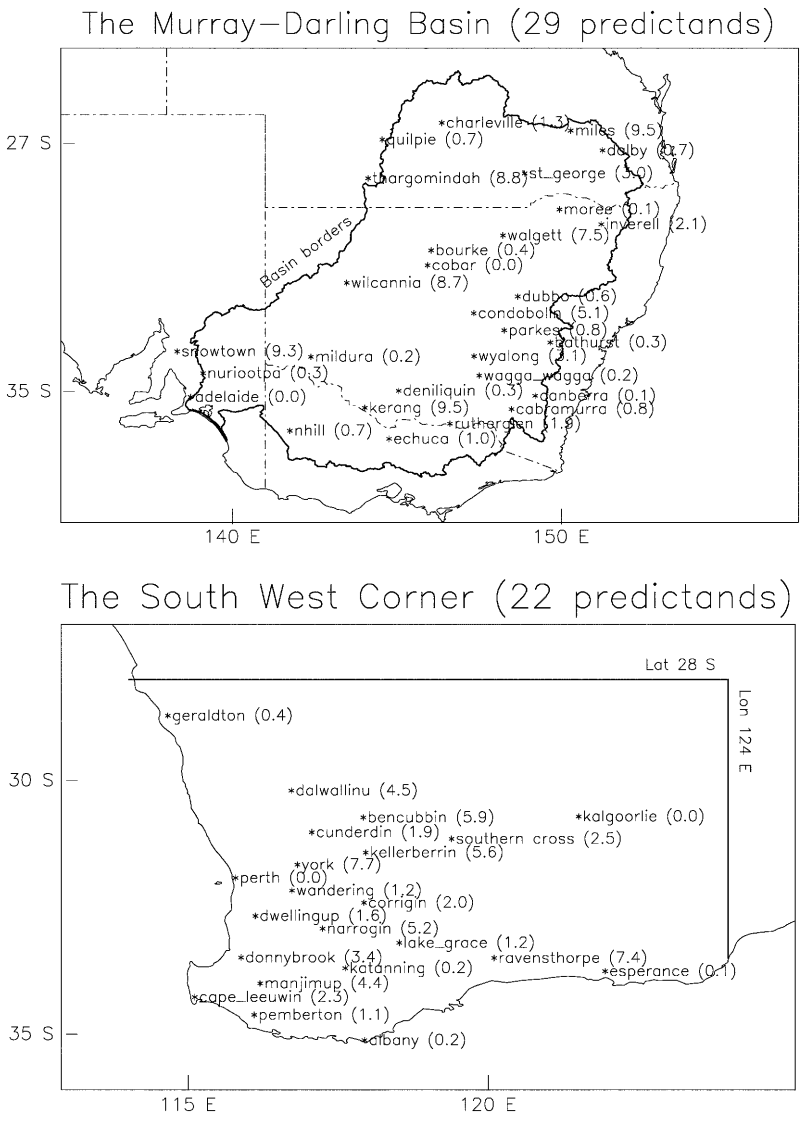

Fig. 1 Locations of temperature stations used. The percentage of missing data is given in *brackets*

production is located. Furthermore, the density of high-quality data is better than for most other parts of Australia. Some additional stations have been added to extend the coverage, they are also of high quality for the entire period of interest in this study 1970–1993.

The reliability of the data is high and very few missing days are reported. Fifty percent of the stations have less than 1% of missing data, however a few stations have up to 10%. To avoid missing data influencing the results, several analogue situations were chosen to back up the best analogue in case synoptic data were missing for this particular date, so that an analogue temperature could always be provided. Results indicate no dependence of the statistical model's skill on the percentage of missing data.

3 Methodology

A range of statistical downscaling methods has been used in the past (Wilby et al. 1998) in regions where sufficiently good local observations are available for model calibration. In this study, the chosen method is the identification of analogues. This technique was first described by Lorenz (1969). Analogues have been used by meteorologists for forecasting as a specific tool or simply as the main driver of forecaster subjective prognosis (Gedzelman 1994). Finally, the experience gained by the Australian Bureau of Meteorology, since 1980, while using analogues for forecasting, has shown the potential of this technique for many variables (Stern 1999).

More recently, analogue techniques have been successfully applied for downscaling in climate studies (Zorita et al. 1995; Martin et al. 1997). They have been shown to be successful at mid-latitudes, producing unbiased series with the right spatial correlation structure. The latter is a very important feature in the Australian landscape where orography is weak and spatial correlation high. Moreover, Zorita et al. (1995) have shown that analogues are capable of reproducing dry and wet spells of precipitation, and in the present study a similar technique is investigated for spells of temperature. More recently, Zorita and von Storch (1999) have shown that, though the use of analogues is a simple technique, it compares well with more elaborate statistical methods.

How to determine the best analogue to describe any particular synoptic situation given by a GCM is the main interest of this study. All predictors are transformed to deviations from the seasonal cycle (seasonal mean is used through the report). To filter the synoptic signal (which has the potential to predict local temperature) from the meteorological noise, the analyses and the model fields are projected onto the leading empirical orthogonal functions (EOFs) or principal components (PCs) (Barnett and Preisendorfer 1978; Sneyers and Goosens 1985). The principal components (PCs)

obtained from analyses are a synthetic way to separate the synoptic signal. In addition to the decomposition into PCs, changes in the geographical extent of predictors was also tested. Reducing the number of points considered produced highly spatially correlated values in direct relationship with local predictants. Optimising the size of the useful domain has proven one of the key issues in selection of the best analogue. The relationship between PC decomposition and domain size is discussed in view of the statistical model results.

The parameters of the statistical model are the set of large-scale atmospheric forcing fields and the metrics to choose analogues. The best analogue of day t is defined as the day t' where the following Euclidean distance is a minimum:

$$d_1^2(t,t') = \sum_{j=1}^{n} [a_j(t') - a_j(t)]^2 \quad \text{with} \quad |t' - t| \leq \Delta t_{cal} \tag{1}$$

When atmospheric forcing fields are decomposed in PCs, a_j refers to PC scores and $n = N_{PC}$ is the number of PCs (usually between 5 and 10): n is chosen to explain a pre-determined level of variance and according to the Craddock and Flood (1969) criterion. PC scores are normalised in order to give the same weight to the variance explained by all PCs. When PCs are not used, a_j are the point values of the atmospheric predictors and $n = N_x \times N_y$ is the spatial dimension.

The parameter Δt_{cal}, defines how close t and t' are in the calendar year. It was tested between 90 days, allowing any date in the same season, and 5 days. When large values are chosen, it is possible to choose as best analogue a day on which, although the synoptic situation was very alike, the actual radiative forcing (due to the solar zenith angle) and other seasonally varying factors were different. This leads to large errors in observed temperature. On the other hand, small values of Δt_{cal} reduce the pool of available situations from which to choose analogues (e.g. 120 cases for $\Delta t_{cal} = 5$). Optimum values of Δt_{cal} were found to vary between 10 and 20 days, allowing 250 to 500 possible analogues.

Due to the limited size of the sampling, no attempt was made to use a weighted combination of the first few best analogues. Although this technique is known to increase the skill of the model (Drosdowsky 1994), it is only suitable if a large pool of data is available, otherwise the standard deviation of the reconstructed series will be strongly reduced.

The first step is to construct, validate and optimise the model. The available dataset is split into two subsets of independent data: 1970–1981 and 1982–1993. The statistical model is applied on the second sample, the temperature series are reconstructed using analogue situations extracted from the first period. The model is adapted for each season and location and parameters are set up during this validation. The main validation tool is the comparison of the reconstructed series with the observed one. Differences of means and variances are compared, as well as the correlations between the two series:

$$Corr = \frac{\sum_{i=1}^{n_{obs}} (T_i - \bar{T})(T_i^a - \overline{T^a})}{\left[\sum_{i=1}^{n_{obs}} (T_i - \bar{T})^2 \times \sum_{i=1}^{n_{obs}} (T_i^a - \overline{T^a})^2\right]^{\frac{1}{2}}} \tag{2}$$

In some cases, the root mean square error (RMSE) between the observed and calculated series was also used:

$$RMSE = \left[\sum_{i=1}^{n_{obs}} \frac{(T_i^a - T_i)^2}{n_{obs}}\right]^{\frac{1}{2}} \tag{3}$$

T is the observed temperature, T^a is the reconstructed one and $\overline{T}$ denotes time averages. The subscript i denotes daily measurement and varies from 1 to n_{obs}.

These statistics are not the only method to measure the skill of a model nor the most complete. However, they were found to be reliable and to provide a clear picture which is easily interpretable in terms of the strengths or weaknesses of the SM. A more synthetic way, the normalised mean square error, as proposed by Williamson (1995), could have been used instead.

In the second step, the optimised model is applied to GCM outputs and the reconstructed temperature series is validated against the same period as the AMIP run: 1979 to 1988. The validation is on the main characteristics of the series: mean, variance and probability density function (PDF).

4 Optimisation of the atmospheric predictors

Several atmospheric predictors were tested, with selection based on these criteria:

1. Being realistically simulated by GCMs
2. Having a strong predictive skill for surface temperature
3. Being complementary to the other predictors

4.1 Validation of the model fields

The comparisons of the GCM fields with the analyses were performed for seasonal means. The overall patterns need to be well represented, although mean biases of the model are removed by using daily anomalies from the seasonal mean. More importantly, the daily variations of the field also need to be well represented.

Mean sea level pressure (MSLP), during Southern Hemisphere winter (June-July-August) over the entire METANAL domain is shown in Fig. 2. Analyses are averaged over the 1979–1988 AMIP period simulated by the GCM. Mean and standard deviations fields are given. The overall pattern is reproduced, with a realistic high-pressure centre over the southeastern part of the Australian continent (overestimated by around 1 hPa) although there is a too-active low belt over the Southern Ocean (by up to 2 hPa). The gradient affecting the south of Australia is too strong. This bias is removed when anomalies are used. The standard deviation given by the model is also realistic; however the model underestimates daily variation from the Equator to 40°S (by less than 1 hPa). Synoptic variations are measured using decomposition into principal components (PCs), to summarise modes of variation. The first five PCs of MSLP in JJA (Fig. 3) given by the model are similar to the ones from the analyses. Furthermore, the percentage of explained variance is of the same magnitude. Overall, this validates the previous hypothesis that the climate (i.e. large-scale features) simulated by the GCM is realistic in terms of broad-scale synoptic variations.

The spatial correlations between the GCM and analyses PCs are summarised in Table 1 for different atmospheric predictors. It appears that dynamical variables (MSLP, Z_{500} and Wind_{500}) are well captured by the model. Spatial correlations of the first five PCs are well above 0.7; except for the third PC of Z_{500}. The percentage of variance explained is very high for MSLP and Z_{500} (over 75%) and similar in the model and the analyses. Less variance is explained by the same number of Wind_{500} PCs, and 4 out of 5 PCs are permuted. Therefore the most reliable representation of the dynamics of the atmosphere is based on MSLP and Z_{500}.

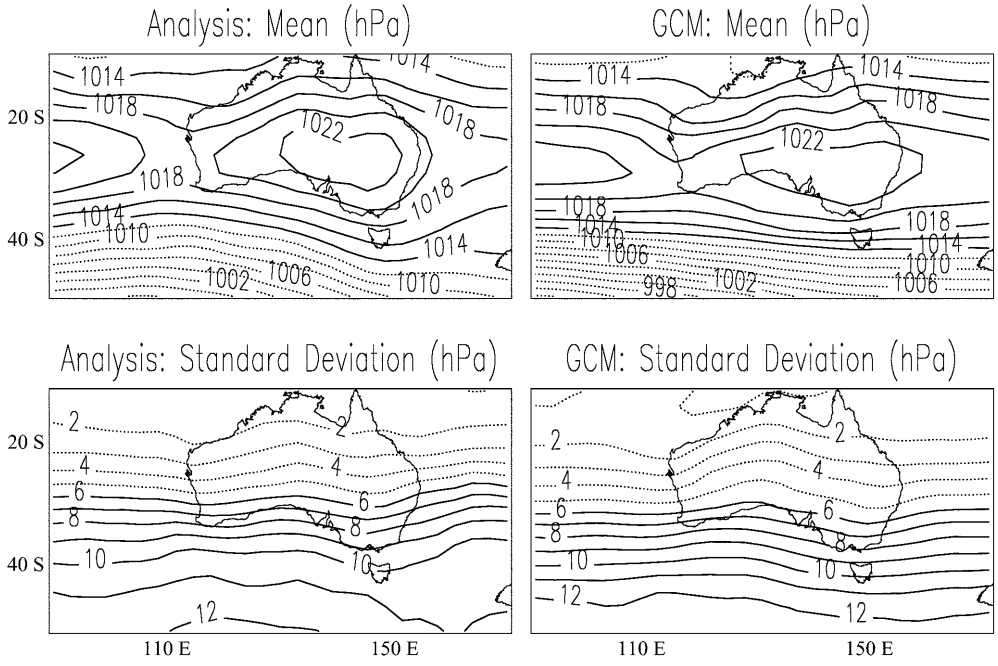

Fig. 2 Mean and standard deviation of MSLP in winter (JJA), averaged over 1979–1988 for METANAL (*left*) and BMRC GCM (*right*). *Dotted lines* are for contours below the aerial average of the field

Next, we analyse variables representing the thermal profile of the atmosphere: T_{850} and the thickness between 1000 hPa and 500 hPa ($Th_{1000-500}$). These PCs tend to be captured less well by the model – apart from the first two – and the order is often mixed. The percentage of explained variance is also smaller.

Finally, the components of the hydrological cycle (RH at 850 hPa and precipitable water) are compared. The correlation between the PCs is marginal and the explained variance by the PCs from the model is very low. Further analyses of the PCs revealed that a large part of the difference is due to temporal non-homogeneities in METANAL. These are due to changes to the analysis system, in particular a strong discontinuity which appeared in 1986 (Fig. 4) when a new regional forecast model was used to assimilate the observations (P. Stewart personal communication). Differences of over 50% are observed in the tropics for the ratio of the difference of the mean of precipitable water before and after 1986 divided by the mean prior to the discontinuity. At mid-latitudes, errors are smaller but show a land-sea contrast which would strongly affect the identification of a suitable analogue. The moisture PCs obtained from METANAL are strongly influenced by this feature and therefore not comparable with the GCM outputs. Furthermore, the large difference between patterns would limit the choice of analogue before or after the discontinuity. For this reason, these variables were abandoned in further development of the SM.

4.2 Predictor skills

The predictive skill of the atmospheric variables has been assessed for each pre-selected field using the correlation between the reconstructed series and the observed ones. Results are shown for winter, using the six leading PCs of each field and averaged over all SWC stations (Table 2).

MSLP is the most skillful predictor, followed by T_{850}, Z_{500} and $Th_{1000-500}$. Wind_{500} is the least skillful of the predictors considered; this is possibly due to the lesser variance explained by the first six PCs. The local time correlations between each atmospheric field and its analogue were calculated. This coefficient measures how close each situation is to its analogue; this is a measurement of the efficiency of using PCs in order to recognise analogue days.

The spatially averaged correlation is around 0.5 for most predictors except Wind_{500} (0.36). The spatial patterns (Fig. 5) show that correlations are a maximum where variability is the highest, as expected, since PC analysis was performed on the covariance matrix. In the case of dynamical fields, the maximum occurs over the Southern Ocean while for thermal variables it extends over the Australian continent.

The Euclidean distance between each day and its analogue has been calculated using Eq. 1 and normalised by the number of PCs. This is a measure of the difficulty of finding suitable analogues: the larger the distance, the more data is needed to identify a suitable analogue. This distance is of the same order for all predictors and does not reveal any weaknesses amongst the predictors considered.

A combination of predictors was also tested. In all cases but one, two predictors gave better predictive skill than one predictor alone. The exceptional case was MSLP and Z_{500}. At mid-latitudes these two variables are strongly related due to the equivalent barotropic character of the atmosphere. The PC decomposition of both variables (not shown) gave very similar results. There-

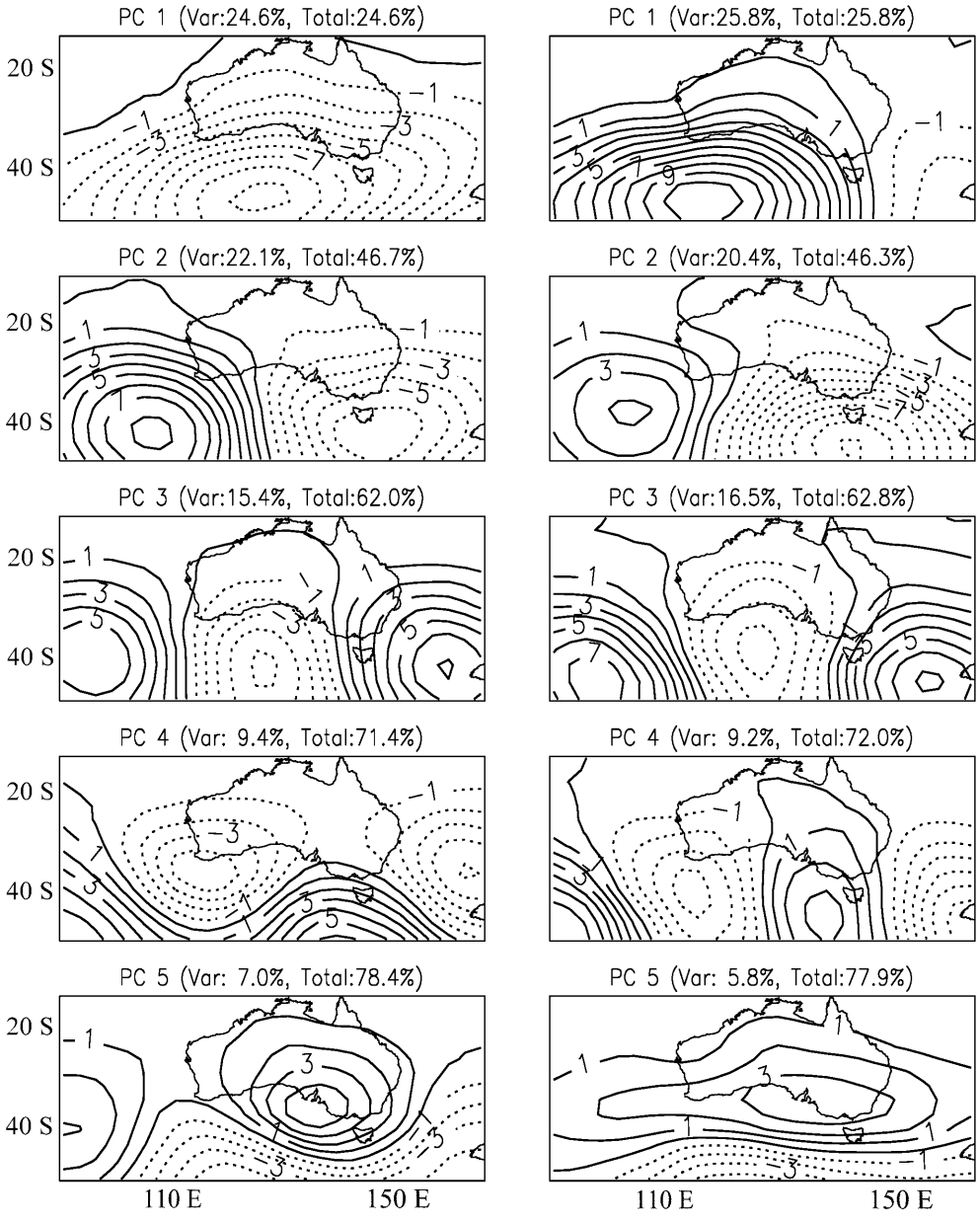

Fig. 3 First five principal components of MSLP in JJA for METANAL (*left*) and BMRC GCM (*right*)

Table 1 Spatial correlation between the PCs obtained from the METANAL analyses and their counterpart from the GCM fields

Predictors	MSLP	Z_{500}	Wind_{500}	T_{850}	$Th_{1000-500}$	RH_{850}	Precipitable water
PC1	0.78	0.96	0.85^{*2}	0.89	0.94	0.43^{*7}	0.77^{*2}
PC2	0.86	0.84	0.74^{*1}	0.93	0.90^{*3}	0.64^{*8}	0.51^{*3}
PC3	0.94	0.68	0.91^{*4}	0.54^{*6}	0.51^{*2}	0.45^{*1}	0.62^{*1}
PC4	0.76	0.75	0.78^{*3}	0.86^{*5}	0.82	0.42^{*10}	0.64
PC5	0.79	0.79	0.90	0.88^{*4}	0.63^{*6}	0.41^{*4}	0.60^{*7}
Variances							
Analyses	78.4	75.7	46.0	58.9	64.1	68.3	53
GCM	77.9	78.1	42.1	51.8	62.1	25.6	35.5

The total variances explained by the first five PCs in both cases is given. The asterisks denote PCs deduced from GCM fields which are ranked in a different order, in terms of explained variance

fore no additional synoptic signal is added when the two fields are combined. Similarly, the combination of T_{850} and $Th_{1000-500}$ does not provide much improvement in skill. These variables describe the thermal profile of the atmosphere and are too similar to be complementary. In general, useful combinations are of a dynamic and a

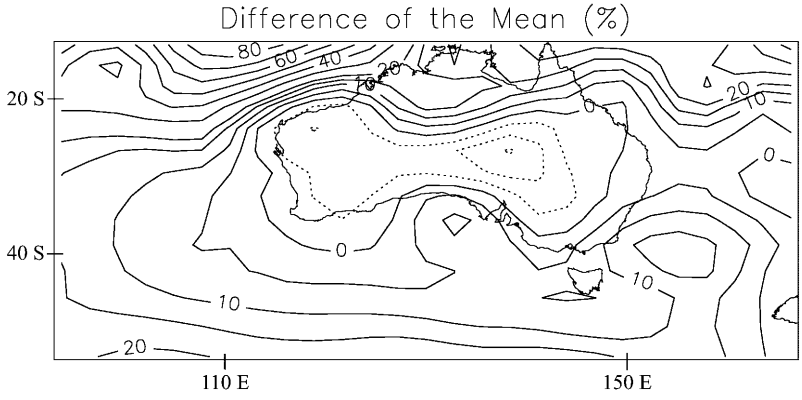

Fig. 4 Difference between precipitable water, in summer (DJF), before and after the 1986 discontinuity expressed as a ratio of the difference over the mean prior to the discontinuity. *Dashed isolines* imply a decrease

Table 2 Skill scores, defined as correlations of the T_{max} series, obtained for various single predictors and for combinations

Predictor(s)	Correlation of T_{max}	Correlation of analogues	Euclid d
MSLP	0.36	0.57	0.44
Z_{500}	0.32	0.48	0.43
Wind_{500}	0.26	0.36	0.43
T_{850}	0.33	0.47	0.42
$Th_{1000-500}$	0.32	0.50	0.42
MSLP and Z_{500}	0.34	0.55/0.47	0.51
MSLP and T_{850}	0.41	0.55/0.45	0.61
MSLP and $Th_{1000-500}$	0.40	0.54/0.48	0.61
Z_{500} and T_{850}	0.39	0.49/0.52	0.50
Z_{500} and $Th_{1000-500}$	0.41	0.51/0.50	0.57
T_{850} and $Th_{1000-500}$	0.34	0.49/0.52	0.50
MSLP, T_{850} and $Th_{1000-500}$	0.42	0.52/0.48/0.50	0.63

Results are shown for winter (June, July, August), using the six leading PCs of each field and averaged over all SWC stations. The second column gives the spatial correlation for the predictors averaged over the entire series between any date and its analogue. The third column is a measure of the normalized Euclidean distance between any day and its analogue, search for in the leading six EOFs dimensional space

thermal predictor together. For instance, MSLP and T_{850} represents the most skillful combination. When combined with $Th_{1000-500}$ the improvement is not significant. Moreover, as analogue situations become less alike, correlation decreases, and the normalised Euclidean distance increases. Therefore, the more predictors used, the larger the pool of historical analyses which must be used to find an equally good match, this is consistent with earlier studies (Van den Dool 1994). Due to the limited dataset available, and as the extra skill added is not significant, MSLP combined with T_{850} is the chosen combination.

4.3 Selection of the domain of interest

Local temperatures are influenced not only by the atmosphere in the vicinity of the measurement but also from further away, due to the spatial correlation of these fields. The domain size applied to predictors to optimise the recognition of synoptic systems from unnecessary background noise is a key parameter of the SM. Different domain sizes (Fig. 6) were tested, ranging from a minimal size just encompassing the domain of interest to the entire METANAL grid. RMSE of the reconstructed series, using two predictors (MSLP and T_{850}), versus the observed one are presented for winter June, July, August (JJA) over the SWC of Western Australia (Table 3). Both primary fields and their decomposition were used. In the latter case the number of PCs used varies for each domain. This number decreases with the size of the domain as more variance is explained by fewer leading PCs. The small domain (not much larger than the area of interest) was found to be the best in most situations and gives better results than analogues based on persistence. Any domain gives better results than a random choice of analogue.

Over the SWC of Australia, an obvious choice would be an off-centre domain extending over the ocean to the southwest of the continent, where prevailing synoptic systems originate. This domain was not optimal, even though this area is important for forecasting weather and, in particular, temperature over the SWC of Australia. The analogues are used as a real time tool to extrapolate local temperature from a coarse grid. In this case the local pattern is more important than that further away, despite this being where the synoptic situation originated from.

An important aspect of this comparison is the decreasing effectiveness of the PCs with decreasing domain size. In large domains, PCs, by filtering synoptic signal from remanent noise, increase the predictive skill from the raw fields. In smaller domains, however, primary variables lead to better results. This feature can be linked to a common criticism made of PCs: they are tightly constrained by their orthogonal nature and the boundary of the geographical domain (Jolliffe 1989). The use of PCs was not pursued further in this study and instead primary fields were used over the small domains.

4.4 Selection of the best metric

The simple metric (method 1) defined previously (Eq. 1) assumes that the current state of the weather system defines the local temperature. Another approach to

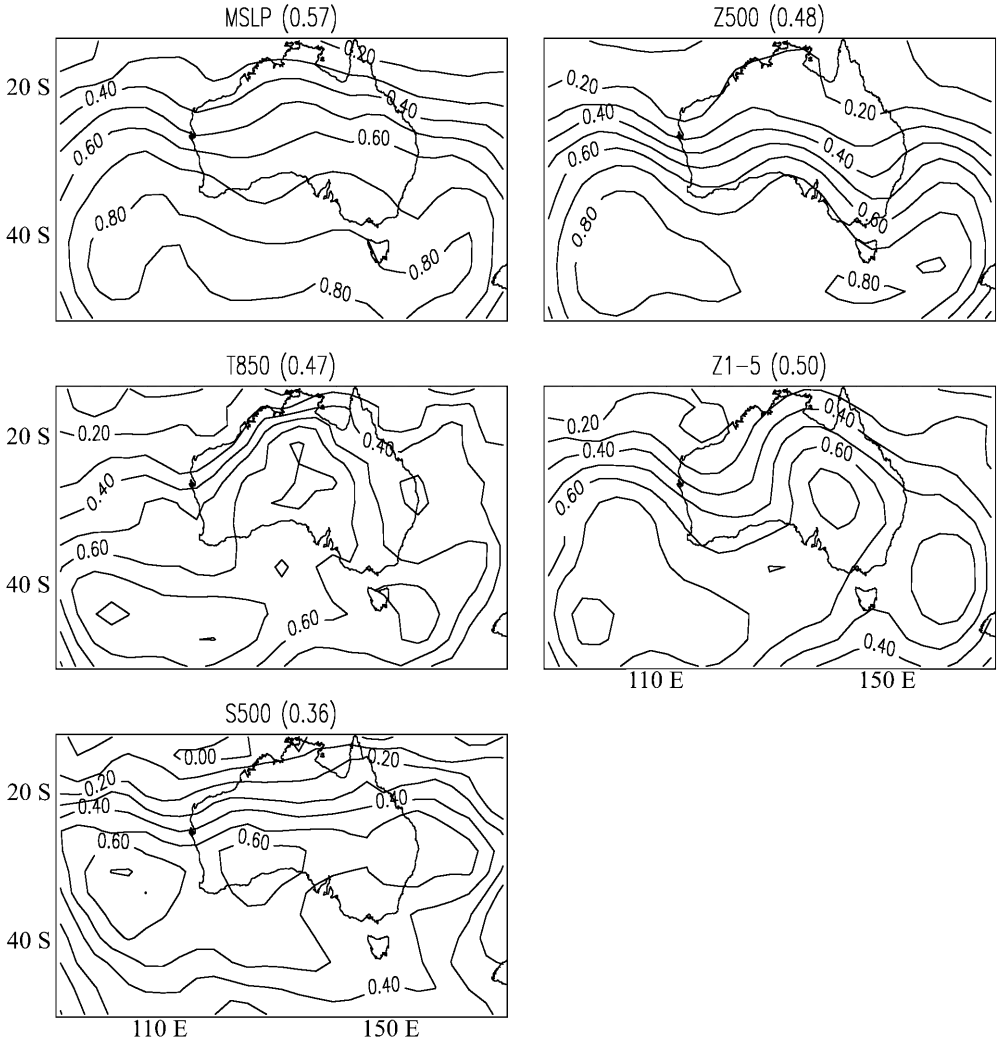

Fig. 5 Time-correlation patterns in winter (JJA) between daily circulations and their analogues for MSLP, Z_{500}, T_{850}, $Th_{1000-500}$ and Wind_{500}. The spatial average of the correlation is given in *brackets*

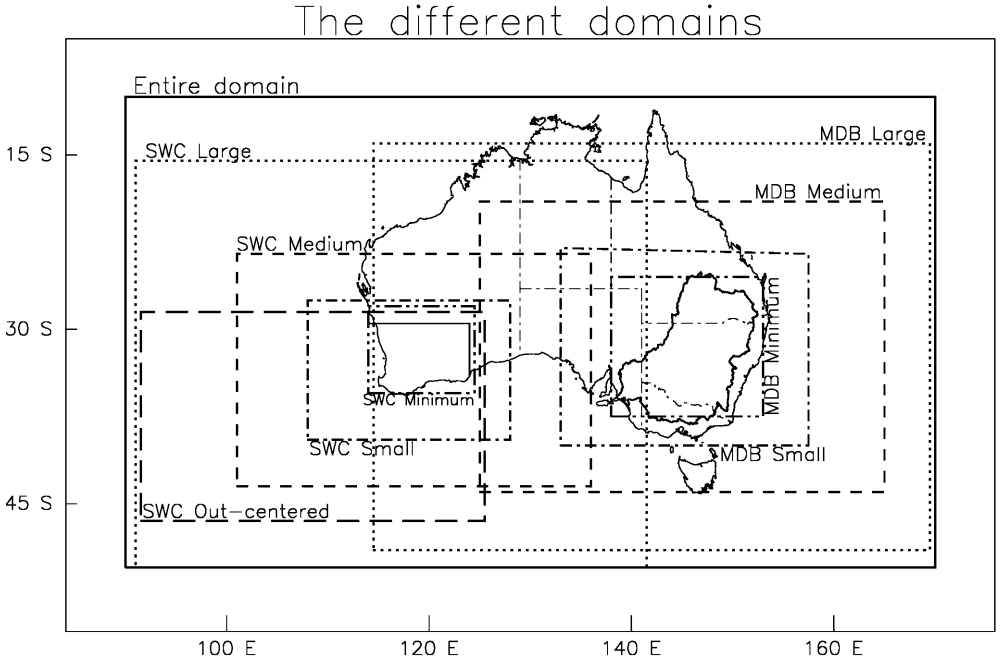

Fig. 6 Domains used for atmospheric predictors

Table 3 RMSE for T_{max} and T_{min} obtained for different geographical domains using either atmospheric predictor raw fields or decomposed into PCs for JJA. Bold figures denote the smallest values

RMSE		Entire	Large	Medium	Small	Minimum	Random	Persistence
T_{max}	Raw field	2.64	2.52	2.33	**2.23**	2.32	3.56	2.41
	PCs	2.57	2.48	**2.34**	2.37	2.49	3.56	2.41
T_{min}	Raw field	3.42	3.26	3.04	**2.90**	2.96	4.12	2.94
	PCs	3.39	3.22	3.14	**3.03**	3.13	4.12	2.94

choose the best analogue, proposed by Barnett and Preisendorfer (1978) would be to take into account the weather trajectory. Therefore a second type of metric (method 2) was defined to take into account the evolution for N days prior to day t:

$$d_2^2(t,t') = \sum_{j=1}^{n}\left[[a_j(t') - a_j(t)]^2 + \sum_{i=1}^{N}[a_j(t'-i) - a_j(t-i)]^2\right] \quad (4)$$

The minimum of this Euclidean distance defines the analogue as the particular day t' for which the evolution of the synoptic situation for N days was the closest to the actual evolution N days prior to day t. It would have been possible to combine the lagged analogues with domain sizes by using larger domains and moving further back in time. The results obtained, however, did not make it necessary to add this extra refinement. Meanwhile, an extension of this method (method 3) with time derivatives of the a_j coefficients was tested but did not prove as successful.

Finally method 2 was extended in method 4 where both the evolution before and after the day t are considered:

$$d_4^2(t,t') = \sum_{j=1}^{n}\left[[a_j(t') - a_j(t)]^2 + \sum_{i=1}^{N_1}[a_j(t'-i) - a_j(t-i)]^2 + \sum_{i=1}^{N_2}[a_j(t'+i) - a_j(t+i)]^2\right] \quad (5)$$

In this method, the analogue is chosen on N_1 days prior and N_2 after the day t. In this case, it is not only the build up of the synoptic situation which is important but also the way the system further evolves.

Methods 2 and 4 were tested for N equal to one or two. Method 2 proved useful in particular for T_{min}, because the minimum temperature is often reached early in the morning around 6 a.m. Local time (LT) while the analyses are for 00*UTC*, between 8 and 11 a.m. LT of the same day. Therefore, taking into account the analysis at day $t-1$ improves the skill of the model by covering the actual time of T_{min} and the previous day.

5 Validation of the statistical model

5.1 Evaluation of the analogue-based temperature series

The results obtained with the SM are presented as seasonal means, averaged over all stations for the MDB and SWC. The SM was optimised using results from the previous section: raw MSLP and T_{850} fields are used over a small domain.

Table 4 Difference in °C between the mean of the reconstructed temperature series and the observed (1982–1993) one. Results are given for seasonal means and the two areas of interest

$\overline{T_{ref}} - \overline{T_{ana}}$	DJF	MAM	JJA	SON
T_{max}				
MDB	−0.04	+0.08	−0.35	+0.35
SWC	−0.71	−0.31	−0.07	+0.17
T_{min}				
MDB	−0.35	+0.61	+0.52	−0.10
SWC	−0.25	+0.31	+0.19	+0.45

Table 5 As in previous table but for differences of variance (in °C^2)

Var T_{ref} − Var T_{ana}	DJF	MAM	JJA	SON
T_{max}				
MDB	+0.85	+2.04	+0.57	+2.99
SWC	+1.14	−0.37	+0.41	+0.52
T_{min}				
MDB	+0.12	+0.41	+1.13	−0.05
SWC	+0.22	+1.18	+1.58	−0.28

The mean of the reconstructed series is compared with the mean of the observed one in Table 4. Biases are small, global values are averaged from contrasting individual results. Furthermore, results are given for the best skill score obtained (with method 2 for T_{min} and method 1 for T_{max} and with Δt_{cal} equal to 10 or 20) and this varies greatly from one set of parameters to another. There is no particular tendency toward positive or negative bias.

Variances are compared in Table 5. Although the magnitudes of the errors are small compared to the variance of the original series, reconstructed series using analogues underestimate the variance of the observed series (in all cases but three). This is most apparent for T_{max} in the MDB. In contrast to the small bias in the mean, this underestimation is coherent from one station to another and robust to parameter changes in the SM. Although results are presented for 1982–1993 as the validation period, similar results were obtained when tested and validation periods were reversed. It is worth noting that since a random re-sampling of the original series would give no biases in mean and variance, this does not infer the skill of the technique but it is a crucial test to insure that the method is not biased.

No clear cut cause is apparent for this bias, in the selection of analogues, toward less extreme cases. It is likely that analogues based on a few atmospheric fields are not sufficient to explain the entire variability of the atmosphere, since part of this variability is expressed by other, independent, variables.

The skill of the SM was then assessed by looking at the correlation (Table 6). To better assess the results of the SM, correlations were compared with persistence. In this case the reconstructed series is the observed series shifted by one day. This is a bench mark for validation purposes often used in forecasting, and often displays skill. In the climate domain, however, persistence which uses information valid locally (although delayed by one day) is not applicable, as the previous day's temperature will never be known. The SM was able to surpass persistence in most cases. Best correlations are achieved for the transition seasons spring and autumn. Comparison with persistence shows that a large proportion of this good score is due to the seasonal trend of temperature which increases the variability around the seasonal mean, in particular in the austral autumn. Results are consistently better than persistence in the SWC for both T_{min} and T_{max}. In winter, where daily temperature variations are small, correlations are low and persistence does as well. Results are less convincing compared to persistence in the MDB, in particular in winter. Overall, the skill of the SM compared to persistence persists all-year-around, except for winter in MDB.

Table 6 Correlation between the mean of the reconstructed temperature series and the observed (1982–1993) one. The two values give the SM value (left) and the value for persistence (right)

Correlation	DJF	MAM	JJA	SON
T_{max}				
MDB	0.65/0.64	0.82/0.82	0.60/0.65	0.81/0.71
SWC	0.76/0.56	0.83/0.73	0.59/0.55	0.83/0.66
T_{min}				
MDB	0.62/0.58	0.73/0.76	0.48/0.60	0.75/0.62
SWC	0.66/0.55	0.69/0.67	0.53/0.52	0.68/0.57

In the SWC of Western Australia, in summer (Fig. 7), best results are achieved near and east of Perth, with correlations above 0.8. Results are very homogeneous with a good spatial correlation, as expected, with the analogue technique which conserves the spatial correlation provided by the predictors. On the fringe of the domain, results are poorer, with correlations falling below 0.7 and as low as 0.52 at Cape Leeuwin. The strong oceanic influence generates very low day-to-day variability which analogues fail to describe fully. The other coastal locations (Perth excepted) obtain low correlation. A possible reason for the limited skill of the analogue is that local temperatures are mostly driven by land-sea interactions such as sea-breezes which are local and only partly taken into account by large scale predictors (Connor 1997).

In the MDB (Fig. 8), correlations have a tendency to decrease from the south west, where the best scores (0.8) were achieved, toward the north east (0.5). Analogues based on the synoptic situation fail to explain a large part of the daily variation of T_{max} in the northern part of the basin. Due to the tropical nature of the weather in summer, daily maximum temperatures are strongly dependent on localised conditions such as convection, cloud cover or soil moisture. It is worth noting that, in this area, persistence achieved much better results. This seems to indicate that the local weather of the previous day is a better guide than large scale patterns. This result is consistent with earlier findings (Timbal and Hender-

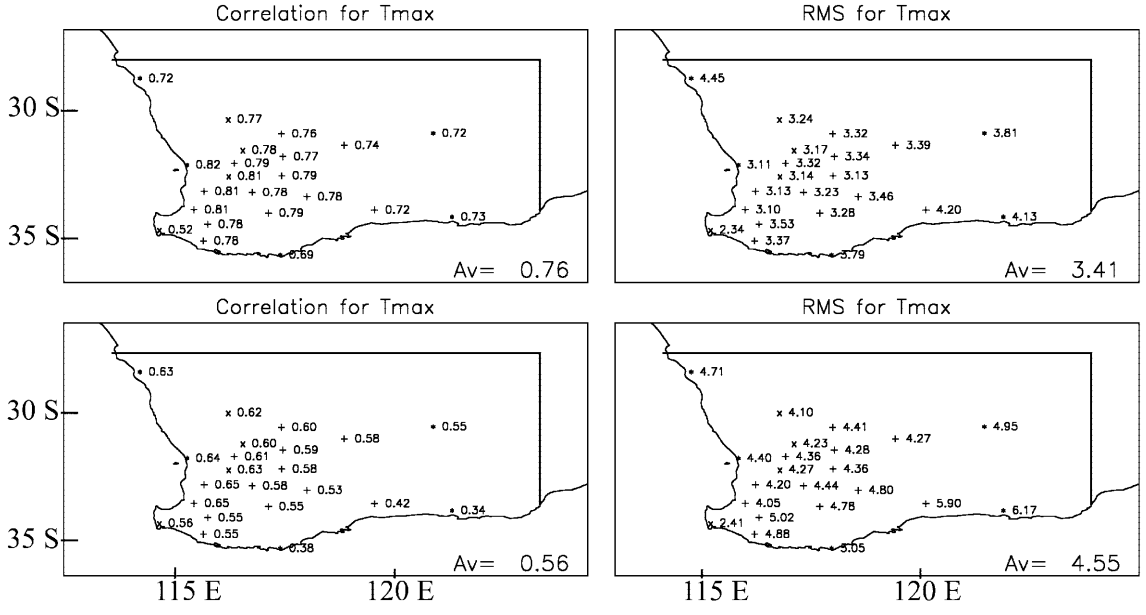

Fig. 7 Correlation for T_{max} in the SWC area given by the statistical model (*top*) and persistence (*bottom*) in summer (DJF)

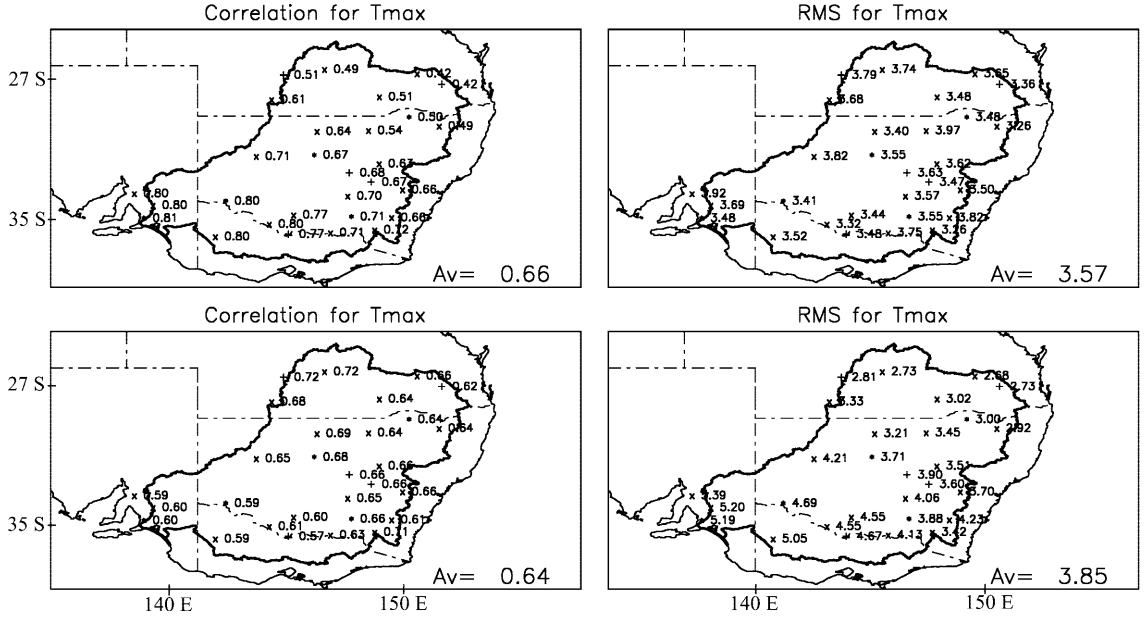

Fig. 8 Same as previous figure but for the MDB

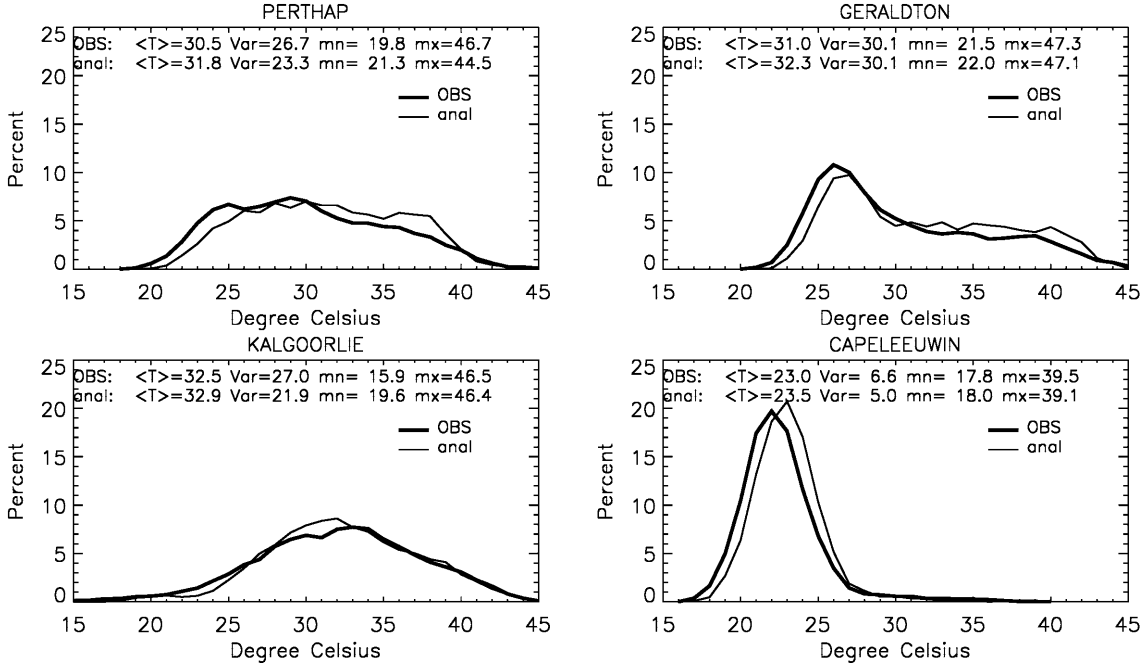

Fig. 9 Probability density function of T_{max}, in summer (DJF), at four stations in the SWC, for 1982–1993 observations (*bold line*) and the reconstructed series using analogues (*fine line*). The mean, variance and extreme values of each series are given

son-Sellers 1998). This suggests that information on the hydrological cycle is needed to provide a more reliable analogue.

5.2 Probability distribution of temperature series

Probability density functions (PDFs) have been plotted at selected stations in the SWC for both the observed and reconstructed T_{max} series (Fig. 9). Four stations have been chosen, as they cover markedly different climate regimes. On one hand, Cape Leeuwin shows a rather constant maximum temperature, around 23 °C, with very few extremely warm days, typical of a coastal site. On the other hand, for inland Kalgoorlie, a broad maximum area between 30 and 35 °C is seen, with symmetrical decreases in probability toward warm and cold days. Geraldton and Perth show a combination of the two previous weather regimes, with a broader area of maximum for Perth, a regular decrease toward warm

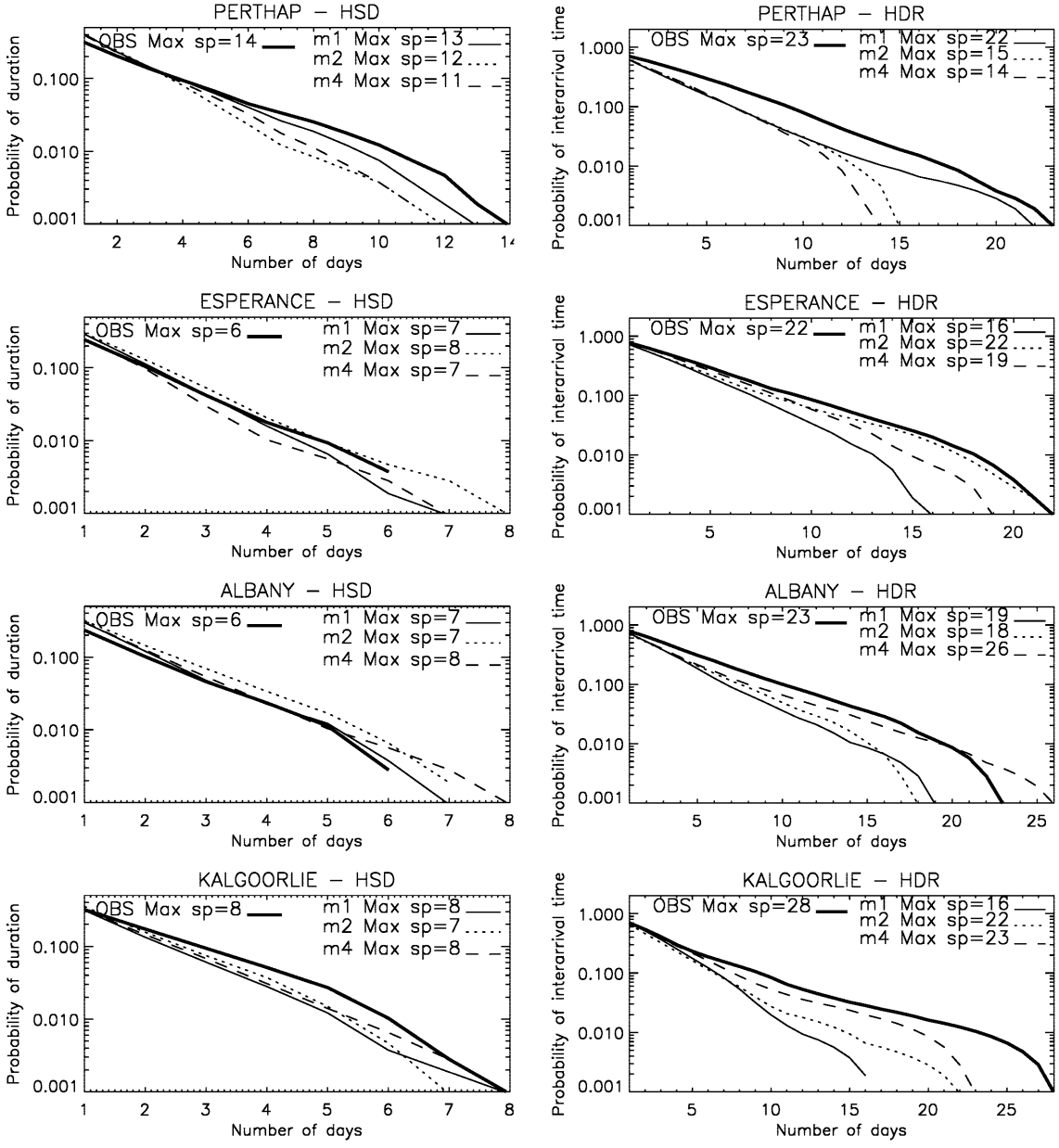

Fig. 10 Hot spell duration (*HSD*) (*left*) and hot day return (*HDR*) (*right*) probabilities, in summer (DJF), obtained with several analogue methods, at four locations in the SWC, compared to 1982–1993 observations

temperatures and a sharp one toward cold days. In all these cases, the reconstructed series match the particular shapes of the individual station PDFs very well. No significant biases are observed for any PDF apart for the maximum probability for Cape Leeuwin being too warm by 0.5 °C.

5.3 Extreme events

One of the major applications of statistical downscaling is to provide daily predictants. The reconstructed daily temperature series allows one to study spells of anomalous temperatures. Such features are particularly crucial when dealing with climate change as these events are the most stressful for human health and farming industries. For example, the impact of heat waves is dependent on their duration. Similarly, long cold spells impact on plant growth (Rosenzweig and Tubuello 1996).

Therefore particular attention has been paid – in the validation of the SM – to assessment of the analogue technique for its ability to reproduce the probability of spell duration. Here, hot spell duration (HSD) is defined as the number of days where:

$$T_{max} \geq \overline{T_{max}} + 0.5 * \sigma_{T_{max}} \tag{6}$$

Similarly, cold spell duration (CSD) is defined as the number of days where:

$$T_{min} \leq \overline{T_{min}} - 0.5 * \sigma_{T_{min}} \tag{7}$$

The mean is calculated from all the days of a particular season and the standard deviation is calculated from daily variations around this mean.

As well as the length of anomalous temperature events, the return period of temperature below or above a particular threshold is also important. In particular, the frequencies of such days have critical impacts on plant growth (Beswick et al. 1998). Both hot day return (HDR) and cold day return (CDR) were defined as the number of days where:

$$T_{max} \leq \overline{T_{max}} + 0.5 * \sigma_{T_{max}} \quad \text{and} \tag{8}$$

$$T_{min} \geq \overline{T_{min}} - 0.5 * \sigma_{T_{min}} \ . \tag{9}$$

In all these cases, frequencies of particular events are plotted as the logarithm of the ratio between the number of events and the total size of the sample (usually around 1000 cases). Therefore, a probability of 1‰ refers to a single case.

It has been shown, that the analogue technique has the ability to reproduce both dry and wet spells (Zorita et al. 1995). But this is a simpler task than for temperature as it is based on a binary outcome (wet or dry days). In the case of temperature, the function is continuous and values are selected according to a threshold. When a spell is broken, it could be by a day which is marginally below or marginally above the threshold.

Both HSD and HDR were plotted for four stations in the SWC (Fig. 10) using the three different metrics defined in Sect. 4. In almost all cases, there is at least one method which matches the observed features well. Where there are discrepancies it is usually toward an underestimation of the length of particular events. There are big differences between methods and no particular method is better, overall. This is worth noting: contrary to spell probabilities, skill-scores were not as sensitive to the choice of method. As these methods differ mainly by the amount of autocorrelation of the temperature series accounted for, it appears that this may be a key component in reproducing anomalous spells.

It is worth noting that the functions are very dependent on the threshold value. If a smaller threshold is imposed for the analogues, the underestimation of the spell duration is strongly reduced. This indicates that even when analogues fail to reproduce a particular spell it is likely that the temperature is still far from the mean but not quite enough to be over the expected threshold.

In the MDB, the probability for HSD and CSD shows interesting contrasts between Mildura and Moree (Fig. 11). Based on the method which provided the best score (method 1 for T_{max} and method 2 for T_{min}) SM results are compared with a random choice of analogues (i.e. using persistence is not a meaningful

Table 7 Biases for each season between the mean of the reconstructed temperature series using the GCM predictor fields and the 1979–1988 observed ones. Each value is compared with bias obtained with GCM-derived temperatures from the nearest grid point (right figure)

$\overline{T_{ref}} - \overline{T_{ana}}$	DJF	MAM	JJA	SON
T_{max}				
MDB	+0.4/+1.8	−0.2/+0.2	−0.1/−0.6	+0.3/−1.1
SWC	−0.7/−1.0	−0.2/−0.9	−0.3/−1.4	+0.1/−2.5
T_{min}				
MDB	−0.1/−3.6	+0.1/−4.2	+0.1/−3.3	+0.2/−4.1
SWC	+0.1/−5.4	+0.1/−4.4	+0.0/−2.1	+0.4/−4.8

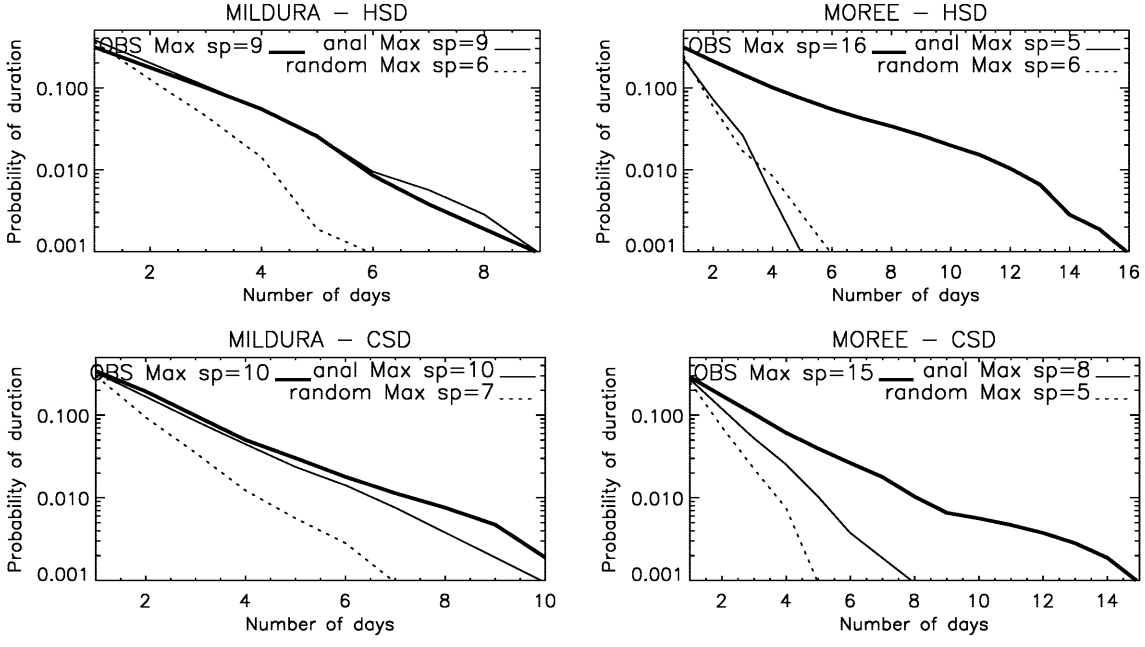

Fig. 11 Cold and hot spell duration probabilities for Mildura and Moree in summer (DJF), obtained with the SM compared with 1982–1993 observations (*bold line*) and a random choice of analogues (*dashed line*)

Table 8 Same as previous table except for differences of variance

Var T_{ref} − Var T_{ana}	DJF	MAM	JJA	SON
T_{max}				
MDB	+1.7/−4.3	+3.2/+3.0	+0.7/+1.1	+5.2/+5.3
SWC	+5.3/+3.3	+6.6/+5.9	+2.1/+1.9	+3.9/+5.4
T_{min}				
MDB	+1.4/−1.7	−1.7/−1.8	+0.0/−0.1	+0.7/+1.6
SWC	+1.7/−1.8	+0.9/−2.7	+0.8/+0.6	+0.0/−1.5

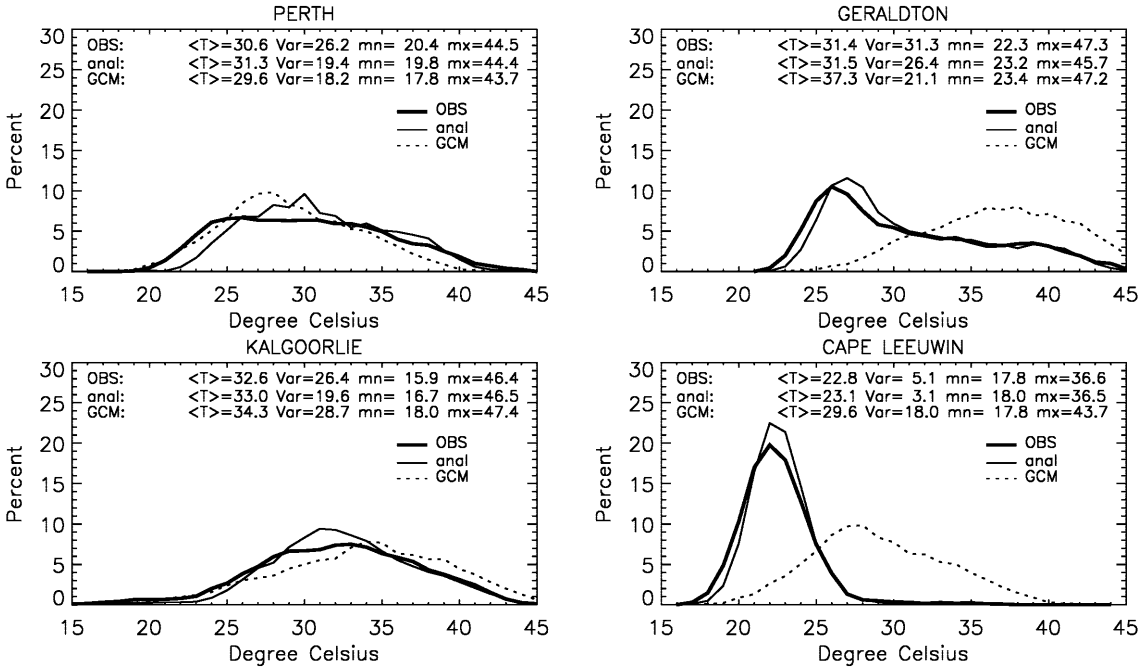

Fig. 12 Probability density function for T_{max} in summer (DJF) from 1979–1988 observations (*bold line*) compared with the reconstructed series using analogues based on the GCM predictors and temperature given by the GCM nearest grid box (*dashed line*)

comparison, as spells match observations perfectly). At Mildura, where the skill-scores were high, the analogues show excellent agreement for both cold and hot anomalous spells. On the contrary, for Moree, where analogues failed to achieve high skill-scores, spells are grossly underestimated. Analogues there provide little better estimation than a simple random choice of the analogues.

5.4 Application to a GCM simulation

This is an important step toward downscaling of climate change scenarios. The optimised SM has been applied to the GCM simulation for each region and the four seasons. In this case, analogues were chosen amongst the available 24 years of analyses (1970–1993) and associated with temperatures observed on the same day. All results are compared with observations during the 1979–1988 period only as simulated by the AMIP-type GCM run. The entire 1970–1993 observations were not used as 1979–1988, with two major *El Niños*, is not representative of the entire 1970–1993 period. To measure the benefit of downscaling GCM large-scale predictors, 2 m-air T_{max} and T_{min} series provided by the nearest GCM grid boxes over land to each station are compared with the analogue-based reconstructed series.

Often GCM surface data shows marked biases when compared with climatology. Therefore, the first check was to determine how well the means of the reconstructed series match observed ones (Table 7). Biases are relatively small and without any particular seasonal or geographical maximum. This was expected from the validation of the methodology which did not show any particular tendency and the comparison of the GCM large-scale fields which were shown to be in agreement with the analyses. In contrast, GCM 2 m temperature shows major biases, up to 4 °C for T_{min}, in both areas and for all seasons.

Similar comparisons were made for variances (Table 8). Reconstructed series using analogues have smaller variances than observed. T_{max}, in particular, is strongly affected, with a reduction of between 0.73 and 6.56 °C. Values are smaller for T_{min}. Variances of GCM predictors have been shown (Sect. 4) to be underestimated compared with the analyses. Although this bias

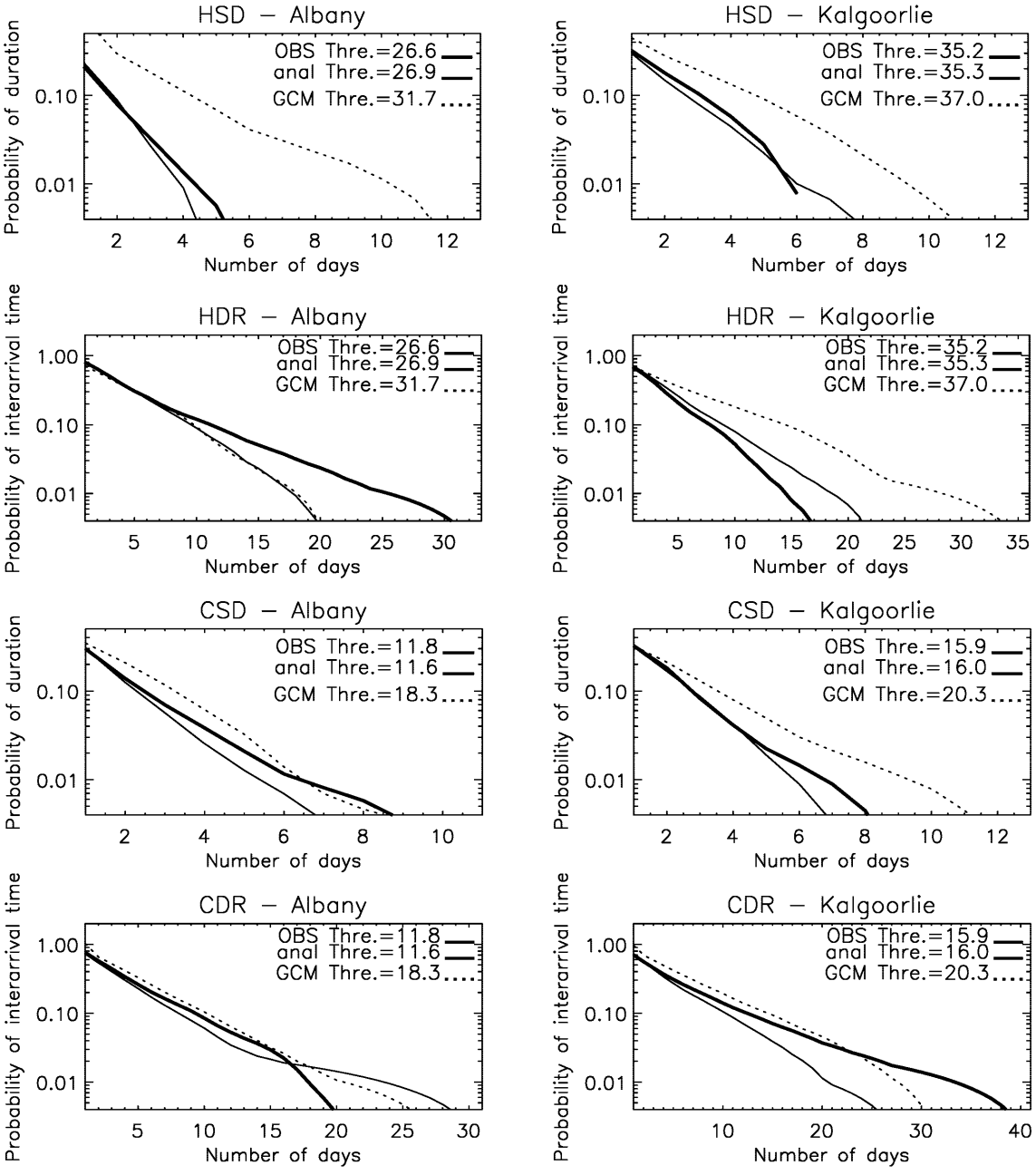

Fig. 13 Cold and hot spell duration and cold and hot day return probabilities for Albany and Kalgoorlie in summer (DJF) as observed (*bold line*) during 1979–1988, for reconstructed analogues using GCM predictors (*fine line*) and direct GCM temperature (*dashed line*)

was small it seems to have major impacts when using the analogue method. In addition, the validation against observations showed a tendency for the method to underestimate natural variability. Therefore the marked reduction is likely to be a combination of those two factors. Results for GCM temperatures show arguably as large differences, but results are contrasted between a reduction of variance for T_{max} and an increase for T_{min}.

The overall quality of the reconstructed series against the observed can be seen in the PDF. In Fig. 12, PDFs are plotted for 1979–1988 observations (bold line), the reconstructed series using analogues and using raw GCM temperature (dashed line), in summer. The same stations were used as during the validation (Fig. 9).

The SM based on GCM predictors successfully reproduces the greatly varying shapes from one station to another. Lower variances can be seen on the PDF. All reconstructed series show differences from observations, in particular toward cold temperatures. Maximum and minimum values, however, are broadly reproduced. The temperatures from the GCM, on the contrary, may be realistic in a few cases (e.g. T_{max} for Kalgoorlie) but in most failed to reproduce particular behaviour, especially

in coastal areas where the GCM fails to reproduce detailed features.

Finally, all types of spells defined previously are examined for two highly contrasting stations: Albany (coastal) and Kalgoorlie (land-locked) (Fig. 13). Probabilities above 4‰ are not shown, as they affect fewer than four cases. The thresholds were adapted to remove the important bias observed in the means for GCM temperatures, as means and standard deviations are taken for each series (rather than being the observed ones). Spells calculated from the GCM T_{max} series are not realistic, while the agreement is better for T_{min} based spells. However, in the latter case, the threshold used is very different from that for observations. By contrast, the agreement between observations and the analogues, in most cases, is good and the thresholds used are realistic. The ability of the analogue model to reproduce analogue spells is maintained when GCM predictors are used.

6 Conclusions and perspectives

A statistical model (SM) has been developed to downscale large-scale predictors for daily maximum and minimum temperature. Such techniques complement dynamical approaches for climate change studies performed with CAOGCM and allow finer temporal and spatial resolutions for impact studies.

The SM is based on the recognition of analogues and was applied for T_{max} and T_{min} series over two agricultural areas of Australia: the MDB and the SWC of the continent. As the SM was optimised, predictors were chosen according to predictive skills and their suitability as GCM outputs. Further tests proved that PC decomposition, though useful to filter the synoptic signal from unwanted noise when the geographical domain is large, was less useful when domains were reduced. The validation of the SM is based on the 24 years of available analyses split into two independent datasets of 12 years each. The validation and optimisation has revealed interesting features of the analogue technique. The reconstructed series shows that the technique is unbiased: the mean of the observation is well reproduced and the probability density functions for almost all stations are quite realistic. However, a general tendency toward smaller variance was noted during the validation, which was further exacerbated when GCM predictors were used.

The SM was found to be more skillful, using correlation, than persistence in all cases except for winter (JJA) in the MDB. Best results were obtained in particular in the austral summer (DJF) and for T_{max}. In the MDB, results are contrasted between the southern part, where analogues are more successful, and the northern part where simulations are not as good. Analogues based on the dynamics of the atmosphere are successful at mid-latitudes, away from tropical influences. Extreme events such as anomalous spells were found to be well captured by the SM, although this shows some dependence on the metrics used.

The main conclusions applied to large-scale fields from a GCM simulation. The ability to reproduce a particular PDF is maintained, with no particular bias on the mean. Furthermore the spells are similarly captured. However the trend toward underestimation of the variance appears more strongly. This indicated that GCMs have enough skill in replicating the PDFS of large-scale predictors.

As expected, the characteristics of the reconstructed series are a major improvement from the GCM-calculated air-surface temperature on the coarse horizontal grid, in particular for the PDFs and the spells. It is clearly demonstrated that statistical downscaling allows a step toward smaller geographical and temporal scales. Some limitations of the technique due to METANAL could be overcome by using the 40 years of re-analyses available from NCEP/NCAR or ECMWF. Of particular interest would be the hydrological variables for both the soil and the atmosphere which were not suitable in METANAL.

Acknowledgements The authors wish to thank several colleagues from the Australian Bureau of Meteorology: R. Dahni for providing the METANAL dataset and guidance with IDL^R, P. Stewart for supplementary analyses and B. Trewin for providing surface observations. Useful comments on the manuscript were made by D. Jones, R. Colman, W. Drosdowsky and D. Jasper. The lead author is supported by the Australian Greenhouse Office.

References

Allan RJ, Haylock MR (1993) Circulation features associated with the winter rainfall decrease in Southwestern Australia. J Clim 6: 1356–1367

Barnett TP, Preisendorfer RW (1978) Multifield analog prediction of short-term climate fluctuations using a climate state vector. J Atmos Sci 35: 1771–1787

Beswick A, Moodie K, Crimp SJ, Jeffrey S, Mullen C, Sweeney S (1998) Climatic considerations for olive orchard site selections using SILO data and climate change scenarios. In: 12th ANZ Climate Forum

Connor GJ (1997) Statistical short-term forecasting technique for wind and temperature in a coastal tropical locations (Townsville). Australian Meteorol Mag 46: 257–265

Craddock JM, Flood CR (1969) Eigenvectors for representing the 500 mb geopotential surface over the Northern Hemisphere. Q J R Meteorol Soc 95: 576–593

Drosdowsky W (1994) Analog (Nonlinear) Forecasts of the Southern Oscillation Index time series. Weather Forecast 9: 78–84

Gedzelman SD (1994) Chaos rules. Weatherwise 47: 21–26

Giorgi F, Mearns LO (1999) Regional climate modeling revisited: an introduction to the special issue. J Geophys Res 104: 6335–6352

Giorgi F, Marinucci MR, Visconti G (1990) Use of limited area model nested in a general circulation model for regional simulations over Europe. J Geophys Res 18 413–18 431

Houghton JT, Filho LG Meira, Callander BA, Harris N, Kattenberg K, Maskell K (1996) Climate change 1995, the science of climate change. Intergovernmental Panel on Climate Change IPCC, University Press, Cambridge, UK

Jolliffe I (1989) Rotation of ill-defined principal components. Appl Statist 38: 139–147

Lorenz EN (1969) Atmospheric predictability as revealed by naturally occurring analogues. J Atmos Sci 26: 636–646

Martin E, Timbal B, Brun E (1997) Downscaling of general circulation model outputs: simulation of the snow climatology of the French Alps and sensitivity to climate change. Clim Dyn 13: 45–56

McAvaney BJ, Colman RA (1993) The BMRC AGCM: AMIP configuration. Tech Rep 38, Bureau of Meteorology Research Centre, Melbourne, Australia

McAvaney BJ, Hess GD (1996) The revised surface fluxes parameterisation in the BMRC AGCM. Tech Rep 56, Bureau of Meteorology Research Centre, Melbourne, Australia

McAvaney BJ, Bourke W, Puri K (1978) A global spectral model for simulation of the general circulation. J Atoms Sci 35: 1557–1583

McAvaney BJ, Dahni RR, Colman RA, Fraser JR, Power SB (1995) The dependence of the climate sensitivity on convective parametrisation: statistical evaluation. Global Planet Change 10: 181–200

Nicholls N (1997) Increased Australian wheat yield due to recent climate trends. Nature 387: 484–485

Palutikof JP, Winkler JA, Goodess CM, Andresen JA (1997) The simulation of daily temperature time series from GCM output Part I: comparison of model data with observations. J Clim 10: 2497–2513

Power S, Tseitkin F, Torok S, Lavery B, Dahni R, McAvaney B (1998) Australian temperature, Australian rainfall and the Southern Oscillation, 1910–1992: coherent variability and recent changes. Australian Meteorol Mag 47: 85–101

Rosenzweig C, Tubuello FN (1996) Effects of changes in maximum and minimum temperatures on wheat yields in the central US: a simulation study. Agric Forest Meteorol 80: 215–230

Seaman RS, Falconer RL, Brown J (1997) Application of a variational blending technique to numerical analysis in the Australian region. Australian Meteorol Mag 25: 3–23

Sneyers R, Goossens C (1985) The principal component analysis: application to climatology and meteorology. Annex to the report on Statistical Methods 9th session of the commission for Climatology, WMO, Geneva, Switzerland

Stephenson DB, Held IM (1993) GCM response of northern winter stationary waves and stormtracks to increasing amounts of carbon dioxide. J Clim 6: 1859–1870

Stern H (1999) Statistically based weather forecast guidance. PhD thesis; available from the Bureau of Meteorology, Australia

Tiedtke M (1989) A comprehensive mass flux scheme for cumulus parameterization in large scale models. Mon Weather Rev 117: 1779–1800

Timbal B, Henderson-Sellers A (1998) Intercomparisons of land-surface parameterisations coupled to a limited area forecast model. Global Planet Change 19: 247–260

Timbal B, Mahfouf J-F, Royer J-F (1995a) Sensitivity of climate change in Europe to the Northern Atlantic warming. In: Proc Int Conf past, present and future climate, SILMU, Helsinki, pp 481–485

Timbal B, Mahfouf J-F, Royer J-F, Cariolle D (1995b) Sensitivity to prescribed changes in sea surface temperature and sea-ice in doubled carbon dioxide experiments. Clim Dyn 12: 1–20

Trewin BC, Trevitt ACF (1996) The development of composite temperature records. Int J Climatol 16: 1227–1242

Van den Dool HM (1994) Searching for analogues, how long must we wait? Tellus 46A: 314–324

von Storch H (1995) Inconsistencies at the interface of climate impact studies and global climate research. Meteorol Z 4: 72–80

Wilby RL, Wigley TM (1997) Downscaling general circulation model output: a review of methods and limitations. Prog Phys Geogr 21: 530–548

Wilby RL, Wigley TM, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS (1998) Statistical downscaling of general circulation model output: a comparison of methods. Water Resources Res 34: 2995–3008

Williamson DL (1995) Skill scores from the AMIP simulations. In: Gates WL (ed) Proc First International AMIP Scientific Conference, WMO/TD-732, pp 253–258

Zorita E, von Storch H (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. J Clim 12: 2474–2489

Zorita E, Hughes JP, Lettemaier DP, von Storch H (1995) Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation. J Clim 8: 1023–1042