Natural Resources and Mines
Managing Queensland's natural resources ... for today and tomorrow

Australian synthetic daily Class A pan evaporation.

Technical Report
December 2005

Queensland Government
Natural Resources and Mines

MANAGING CLIMATE VARIABILITY
R & D PROGRAM

Cooperative Research Centre for
IRRIGATION FUTURES

Australian synthetic daily Class A pan evaporation.

Technical Report
December 2005

David Rayner[1]

1. Queensland Department of Natural Resources and Mines
80 Meiers Rd Indooroopilly
Queensland Australia 4064
http://www.nrm.qld.gov.au/silo/

Project: LWA QNR-37

Funded by:

Queensland Department of Natural Resources & Mines,

Managing Climate Variability Program, Land & Water Australia.

Cooperative Research Centres for Irrigation Futures

Delivered through SILO, a partnership between:

The Australia Bureau of Meteorology, and

Queensland Department of Natural Resources & Mines.

Department of Natural Resources and Mines, Queensland
December 2006

QNRM05435
ISBN 1 74172 003 6

General Citation

Rayner, D.P. (2005), Australian synthetic daily Class A pan evaporation. Queensland Department of Natural Resources and Mines.

General Disclaimer

The information contained in this publication is intended for general use, to assist public knowledge and discussion and to help promote sustainable natural resource management. The information does not represent a statement of the policy of the Department of Natural Resources and Mines, Queensland. The information should not be relied upon for the purposes of a particular matter. Appropriate enquiries must be made to establish the current accuracy of the information and whether any new material is available on the subject matter; legal advice should be obtained before any action or discussion is taken on the basis of any material in this document. The Department of Natural Resources and Mines, Land and Water Australia and their partners, the CRC for Irrigation Futures, and the author assume no liability of any kind whatsoever resulting from any person's use or reliance upon the contents of this document.

Enquiries should be addressed to:

David Rayner
David.Rayner@qld.gov.au
Department of Natural Resources and Mines
80 Meiers Rd Indooroopilly
Queensland Australia 4068

Table of contents

1 Acknowledgements

I am very grateful to David Jones (Australian Government Bureau of Meteorology), Francis Chiew (University of Melbourne), Michael Roderick (Australian National University) and Wayne Meyer (CSIRO) for their detailed and constructive reviews of this report. I also thank Alan Beswick, Keith Moodie, Li Fitzmaurice and Greg Kociuba (Queensland Department of Natural Resources and Mines, NR&M) for their many useful suggestions and tireless maintenance of the SILO system. Finally, I acknowledge the guidance and motivation provided by Greg McKeon (NR&M), who wrote "$e = a + b.R_s + c.vpd$" on my whiteboard in 2002, and consequently started this whole project.

This research was funded by Queensland Department of Natural Resources and Mines, Land & Water Australia through the Managing Climate Variability R&D Program, and the CRC for Irrigation Futures.

The Managing Climate Variability R&D Program is a partnership between Land & Water Australia, Grains Research and Development Corporation, Meat & Livestock Australia, Natural Heritage Trust, Australian Government Department of Agriculture, Forestry and Fisheries, the Rural Industries Research and Development Corporation, Sugar Research and Development Corporation and Dairy Australia. Associate supporters include the wool industry's Land, Water & Wool initiative and the National Farmers' Federation.

2 Summary.

Queensland Department of Natural Resources and Mines (NR&M) has produced a synthetic, daily Class A pan evaporation dataset for Australia for the period 1910–present. This modelled pan evaporation dataset is a more accurate estimate of daily Class A pan evaporation than the long-term averages which are often used for the pre-1970 period. However, the synthetic pan does not represent the largest variations in daily observed pan evaporation, nor does it represent climate trends.

The synthetic pan evaporation model is a simple linear combination of gridded solar radiation and vapour pressure deficit. The model coefficients were calculated with multiple-regression using gridded data for the period 1975–2003. Regression coefficients were derived individually for each pixel in the NR&M climate grids, and independently for each month.

The synthetic pan dataset was produced using the linear model and the NR&M climate grids. These grids cover all Australia at 0.05º (~5 km) spatial resolution. The resulting synthetic pan evaporation grids cover the time period 1910–present, although the accuracy of the grids will be lower in the pre-1957 period.

3 Review.

3.1 Class A Pan evaporation observations.

Evapotranspiration is the transfer of water from the landscape to the atmosphere. Evapotranspiration is a critical component of the water cycle, and an important component of many agricultural, hydrological and climate models.

Because evapotranspiration is difficult to measure directly, atmospheric evaporative demand is often estimated by observing the evaporation rate in an open water container. The Class A pan evaporimeter is the standard instrument used for such observations in Australia. This is a circular pan made of galvanized iron, 4 ft (121 cm) in diameter and 10 in (25 cm) deep, mounted on an open wooden platform. In Australia, the pans are protected with a wire “bird-guard” to stop animals affecting the measurements.

The Australian Government Bureau of Meteorology (ABM) began systematically installing Class A pans in the 1960s, with the universal installation of bird-guards following a few years later. By 1975, the ABM had a nearly homogeneous evaporimeter network[1].

3.2 Interpolated Class A pan evaporation data.

Spatial modelling applications require gridded climate datasets (which are estimates of climate variables over large areas) rather than observed climate records. The Queensland Department of Natural Resources and Mines (NR&M) has created a gridded daily pan evaporation dataset by interpolating the ABM Class A pan evaporation records (Jeffrey et al. 2001). Because early pan evaporation records are considered inconsistent and unreliable (Hounam 1961), pan evaporation records prior to 1970 were not used to generate interpolated surfaces.

Instead, many applications which require spatial pan evaporation data prior to 1970 use grids of seasonally-dependent, long-term average values as a substitute. Using substituted long-term averages imposes limitations on the accuracy of any historical water-balance model.

3.3 Modelled synthetic Class A pan evaporation data.

One solution is to develop *synthetic* Class A pan evaporation grids, where pan evaporation is modelled using other climate variables. An unbiased, universally-applicable synthetic pan evaporation model would be very valuable; a model which operated successfully across different climate regimes and seasons could be applied with confidence to historical and future time-series, where climate changes may be important.

A synthetic pan evaporation model was investigated by Richards et al. (2001), as part of the AussieGRASS project (Carter et al. 2000). For selected Queensland rangeland locations, this model accounted for 70–80% of the variability of measured pan evaporation. This model used multiple-regression of pan evaporation onto observed meteorological variables. The regression used an “ensemble” approach, using data from multiple sites to obtain a single set of regression coefficients. We evaluated this approach and found that systematic regional and seasonal biases in the model residuals made it unacceptable.

Class A pan can be estimated using the “Penpan” formula (Linacre 1994). This is a modified form of Penman’s evaporation equation (Penman 1948) that incorporates such factors as the effect of pan geometry on incident direct solar radiation, diffuse solar radiation incident on the pan wall, and solar radiation reflected off the surrounding ground surface onto the pan wall. The Penpan formula also requires an estimate of daily wind. We evaluated this model using observed daily and monthly climate data and found that, although the bias was much lower than the

[1] For details see the ABM CD-ROM “Australian Daily Evaporation IDCJDC05”.

Richards et al. model described above, the typical error was considerably larger than for the pixel-based regression model discussed below. More importantly, the main advantage of the Linacre model is supposedly its universal applicability, but this would be compromised in our application because there is no reliable, gridded historical wind dataset for Australia.

3.4 A pixel-based regression model for pan evaporation.

We decided to postpone using a physically-based pan evaporation model until suitable gridded historical wind data become available. Instead, the synthetic pan model described in this report is a spatially and seasonally-varying, empirical model. Regression coefficients for pan evaporation onto solar radiation and vapour pressure were determined independently for each pixel of the standard NR&M pan evaporation surfaces, and separately for each month. This eliminates regional and seasonal bias, but the cost is that the model is not guaranteed to be robust against regional climate variability or climate change. However, as illustrated in Section 7, the model is clearly a better estimate of daily Class A pan evaporation than the seasonal averages which are currently used in many modelling applications as surrogates for pre-1970 data.

4 Overview of input data.

The synthetic pan was modelled using NR&M daily climate grids (Jeffrey et al. 2001). These grids were calculated using a thin-plate spline algorithm to interpolate observed data supplied by ABM. However, not all of the climate variables are directly observable. The required climate variables were generated from the observations as follows:

- Minimum and maximum daily temperatures and pan evaporation were recorded directly by observers,
- Vapour pressure records were derived from 9 am wet-bulb/dry-bulb air temperature observations, and station-level air pressure,
- Daily solar radiation records were derived from total cloud amounts. Cloud amounts are observed using standard meteorological procedures, and were supplied in oktas[2].

Details and references for the vapour pressure and solar radiation calculations are provided in Jeffrey et al. (2001).

Vapour pressure deficit grids were calculated using the NR&M maximum and minimum temperature grids to estimate daily-average saturation vapour pressure, and using the NR&M 9am vapour pressure grids to approximate the daily-average actual vapour pressure. Details are given in Section 5.

The NR&M climate grids for the period 1910–1956 were generated using data from a much sparser observing network than was used for the 1957–present period. The grids for this earlier period have higher biases and errors (Rayner et al. 2004). Consequently, the synthetic pan evaporation for 1910–1956 will also have higher biases and errors. This issue is discussed further in Section 9.5.

The observed pan evaporation data supplied by ABM were also used for the validation analyses described in Sections 7, 8 and 9.

4.1 Bird-guards

The pan evaporation dataset supplied by ABM contains Class A pan observations made both with and without bird-guards fitted. Pans without bird-guards are known to have higher evaporation rates (van Dijk 1985), and the observation records may contain errors caused by animals drinking or splashing water out of the pan. Unfortunately, the NR&M evaporation grids currently include observations from pans without bird-guards, without any correction factors applied. This situation is currently being revised.

The ABM CD-ROM does contain a file which lists dates by which stations had bird-guards fitted. This suggests that most stations had bird-guards fitted by 1975, or shortly thereafter. Only five stations[3] recorded a significant amount of data after 1975 without a bird-guard. Because only post-1975 data were used to calibrate the synthetic pan model, the calibration will not be significantly affected by pans without bird-guards.

4.2 Gridded pan evaporation properties.

Australian mean daily pan evaporation, calculated from the NR&M gridded data, is shown for reference in Figure 4.1.

---

[2] An okta is the number of eights of the sky obscured by cloud.

[3] Douglas River (14901), Te Kowai Exp Stn (33047), Mount St Leonard (86142), Geelong Salines (87023), and Scottsdale (91219).

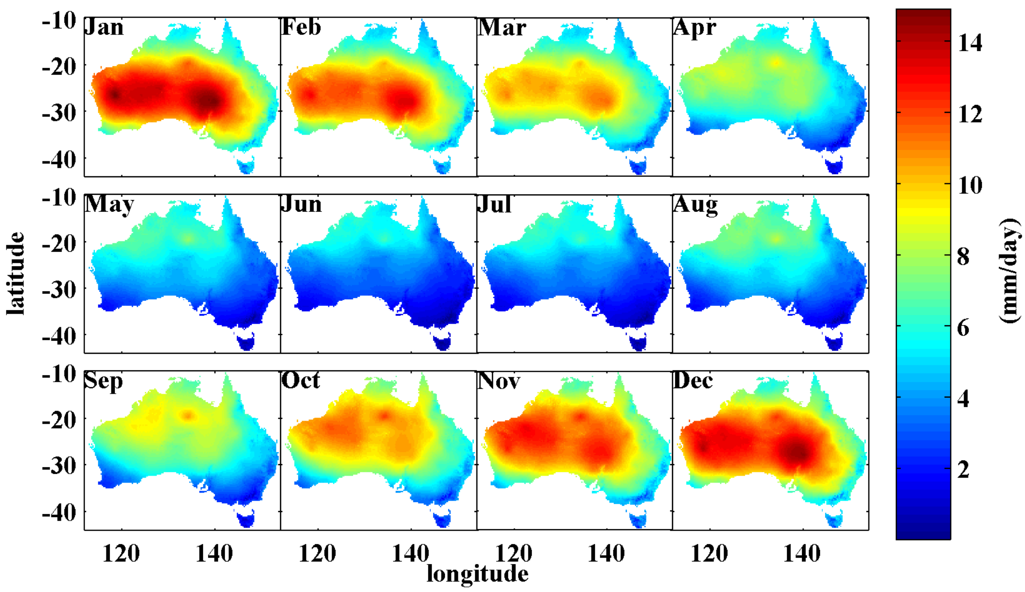

Figure 4.1: Mean daily evaporation for each month, calculated using data for 1975–2004.

Correlations between gridded daily pan evaporation and solar radiation, and between daily pan evaporation and vapour pressure deficit, are shown in Figure 4.2. and Figure 4.3 respectively. Note that the correlation between pan evaporation and vapour pressure deficit is generally higher than the correlation between pan evaporation and solar radiation. The synthetic pan evaporation will be more accurate for regions and months where the correlations in Figure 4.2 and Figure 4.3 are high.

The correlations between daily pan evaporation and solar radiation shown in Figure 4.2 match the inter-day solar radiation variability (not shown). This is presumably because, if the daily solar radiation is low or nearly constant within the month, then variations in other climate parameters will obscure the relationship with evaporation. The same reasoning probably explains the low correlation between daily pan evaporation and vapour pressure deficit (Figure 4.3) in southern Australia from April to September. However, the low correlation north of latitude -20º from May to September is perplexing. Interestingly, this band of low correlation is even more pronounced in the correlation between weekly pan evaporation and vapour pressure deficit.

The small, discrete regions of depressed correlation, most noticeable in Figure 4.3, are all near pan evaporation observing stations. Similar features are also present in de-trended data, which demonstrated that these features are not primarily caused by long-term changes in the relationship between pan evaporation and other climate variables. The features most likely arise because, in regions away from observing stations, interpolation effectively averages together data from several nearby stations. This averaging reduces the contribution from observation errors (see Section 8.2) and unaccounted, site-specific effects.

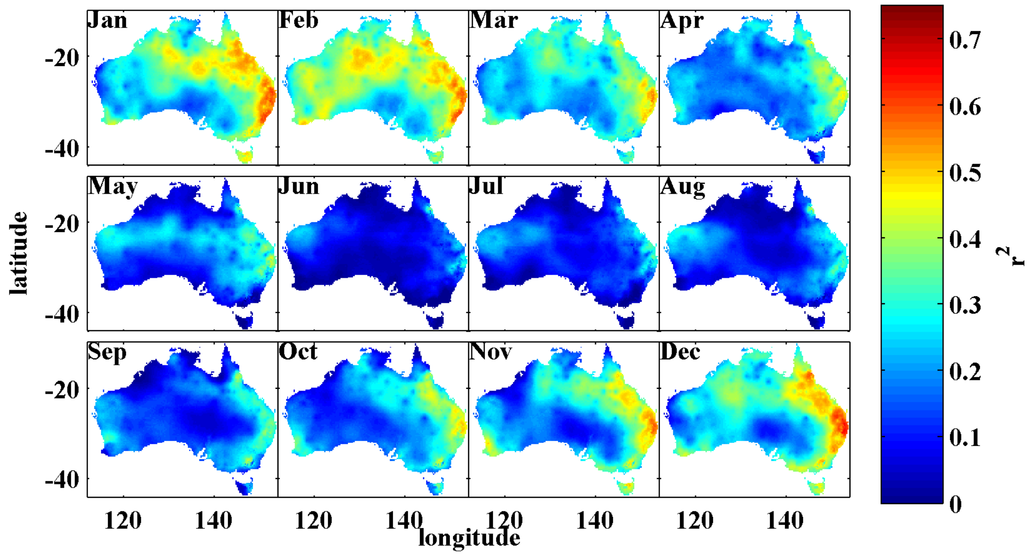

Figure 4.2: The coefficient of determination (r^2) between gridded daily evaporation and gridded daily solar radiation. The correlation is highest for summer months, for eastern Australia.

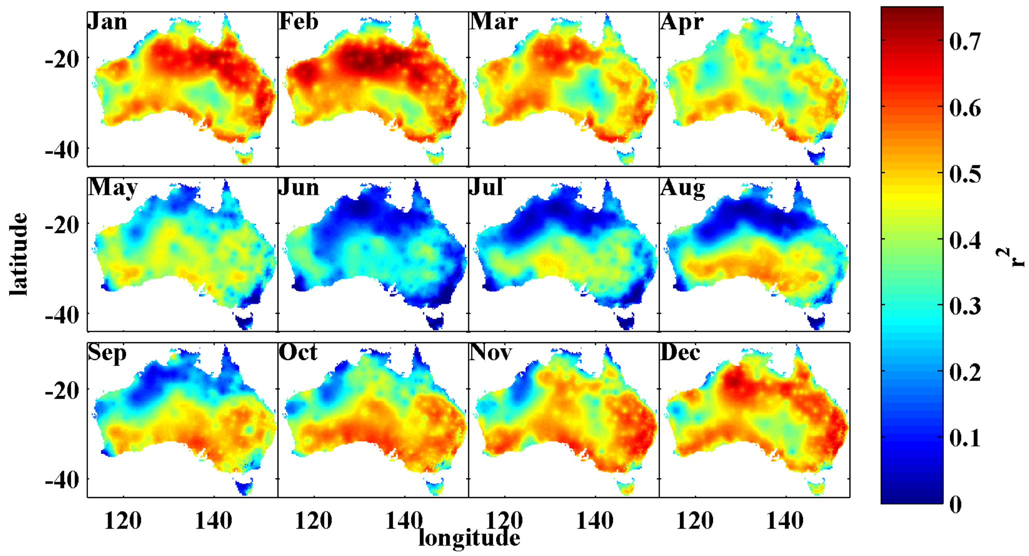

Figure 4.3: The coefficient of determination (r^2) between gridded daily evaporation and gridded daily vapour pressure deficit. The correlation is low over much of northern and south-east Australia during winter months. The discrete areas of depressed correlation are coincident with the pan locations.

5 The pan evaporation regression model.

Gridded Class A Pan evaporation was modelled as:

$$E_{pan} = b \times R_s + c \times (e_s - e_a) \tag{5-1}$$

where:

E_{pan} gridded Class A pan evaporation rate [mm day^{-1}],

R_s gridded daily total solar radiation [MJ m^{-2} day^{-1}],

e_s gridded mean daily saturation vapour pressure [hPa],

e_a gridded 9 am vapour pressure [hPa], and

b, c are empirically-derived coefficients.

The term $(e_s - e_a)$ in Equation 4-1 is known as the vapour pressure deficit.

The mean daily saturation vapour pressure is estimated as (Jeffrey et al. 2001):

$$e_s = 0.75 e^0(T_{max}) + 0.25 e^0(T_{min}) \tag{5-2}$$

where

$e^0(T)$ saturation vapour pressure at air temperature T, given by (Allen et al. 1998):

$$e^0(T) = 6.108 \exp\left(\frac{17.27T}{T + 237.3}\right) \tag{5-3}$$

Note that, because saturation vapour pressure is a nearly linear function of temperature over typical Australian temperature ranges (r^2 between saturation vapour pressure and temperature is 0.98 over the range 10°C–30°C), to first approximation changing the coefficients in Equation 5-2 only rescales the coefficients in Equation 5-1, it does not improve the accuracy of the model.

Using Equation 5-2 to estimate daily-average saturation vapour pressure, and using 9am vapour pressure to estimate daily-average vapour pressure, can return zero or negative vapour pressure deficits. For the purpose of this analysis, negative vapour pressure deficits were treated as zero.

Physically, evaporation is actually related to net radiation rather than solar radiation. A physical model for the net radiation received by a Class A evaporation pan is presented by Linacre (1994). However, to first order this conversion is handled implicitly by the regression algorithm, and by solving for the regression parameters independently for each month and grid-pixel.

The parameters *b* and *c* in Equation 5-1 were fitted for each pixel in the 0.05º grid using multiple regression, using daily data from 1975–2003 inclusive. Parameters were fitted separately for each month. The spatial and seasonal variability of the parameters *b* and *c* is shown in

Appendix A. The spatial and seasonal variability of the synthetic pan and the model components is shown in Appendix B.

Visual inspection of residuals did not suggest any correlation with other available NR&M gridded climate variables. In particular, despite the association of rainfall with synthetic pan anomalies discussed in Sections 7.1 and 8.1, rainfall does not systematically explain a significant component of the synthetic pan residuals.

5.1 Comments on the regression model.

The parameters *b* and *c*, and their spatial and seasonal variability as shown in Appendix A, do not necessarily have any physical significance. Because solar radiation and vapour-pressure deficit correlate with each other as well as with pan evaporation, spatial variation of the parameters *b* and *c* may simply represent the mutually-correlated pan evaporation being transferred between the two terms.

Changes in the heat energy of the pan water are not included in the synthetic pan model. Intuitively, changes in the heat stored in the pan are expected to dampen the daily evaporation time-series response to changes in radiation and vapour-pressure deficit. As discussed in Sections 7.1 and 9.2, however, it is actually the synthetic pan time-series which is damped compared to the observed evaporation time-series. This suggests that heat storage in the pan is not as important as the other errors discussed in Section 9.2.

To get a perfect pan evaporation model, site and instrument conditions together with local[4] topography and weather would need to be incorporated. For the moment, such factors are simply treated as observational error, and are ignored. Local factors should not cause regional biases between the modelled pan evaporation and the measured pan evaporation.

Regional biases, on the other hand, mean that the modelled pan evaporation is systematically misrepresenting the relationship between the observed climate parameters and evaporation. The empirical, pixel-based, monthly regression model largely eliminates regional and seasonal bias, but the cost is that the model is not guaranteed to be robust against regional climate variability or climate change (see Section 9.4).

An alternative approach to modelling synthetic pan evaporation would have been to model the pan evaporation using observed data at each site independently and then interpolate the resulting site-based evaporation using a thin-plate spline. Conceptually, this technique has an advantage in that it could be calibrated using observed, rather than interpolated, climate data. In practice, however, this would use the observed daily climate data very inefficiently, because the majority of daily climate observing stations do not have an evaporation pan. Further, only about half of the observed pan evaporation records have simultaneous measurements of all the other required climate variables. In contrast, the modelling approach used here makes use of all available daily climate records, through their effect on the daily climate surfaces.

---

[4] In this context, “local” means “at a scale less than the pixel size”.

6 Properties of the synthetic pan evaporation data.

The synthetic pan evaporation model estimates daily Class A pan evaporation better than do the seasonal averages which are currently used as surrogates for pre-1970 data. The synthetic pan evaporation model:

- has some representation of observed daily pan evaporation variability, as shown by comparison against observed time-series (Section 7.1),
- has improved overall accuracy (Section 7.2), and
- preserves the natural inter-correlations within the climate dataset, especially the relationship with rainfall (Section 7.3).

Interestingly, the synthetic pan evaporation model is also free of some of the errors which characterise observed pan evaporation. In particular:

- pan evaporation records contain unrealistically high values, often associated with high rainfall (Section 8.1), and
- many pan evaporation observations are rounded to the nearest 4 mm (Section 8.2).

However, the synthetic pan model does not represent many important properties of the observed pan evaporation record.

- The model does not use wind, which clearly has an important effect on pan evaporation (Section 9.1).
- Perhaps as a consequence, the synthetic pan evaporation does not represent days with very high evaporation (Section 9.2).
- The synthetic pan model does not replicate trends in observed pan evaporation (Section 9.3).
- Because the model is spatially and seasonally dependent, it will not be robust against climate variability and climate change (Section 9.4).

6.1.1 The observed pan evaporation validation dataset.

Many of the analyses described in this section compared synthetic pan evaporation data with a validation dataset of observed pan evaporation data. This comparison revealed features in both the observed and synthetic data.

The validation data were sourced from the ABM CD-ROM "Australian Daily Evaporation IDCJDC05", the same data used to generate the NR&M gridded evaporation data. For the validation data, only observing stations which had pan evaporation records for at least 50% of days in the period 1975–2003 were used. This selection process yielded 1.6 million station-days of data. The distribution of the observing stations used for validation can be seen, for example, in Figure 7.2. Note that the validation data are not evenly distributed spatially, and inland areas with high evaporation rates are under-represented.

Although no cross-validation was performed, I emphasise that the validation data consisted of observed pan evaporation data only; gridded pan evaporation data were not used for validation.

7 Synthetic pan evaporation improvements.

This section demonstrates that the synthetic pan dataset is an improvement over the use of long-term average pan evaporation data.

7.1 Improved representation of daily pan evaporation variability.

The example time-series shown in Figure 7.1 shows that the synthetic pan evaporation does represent, to some extent, the daily variability in the observed pan evaporation. The synthetic pan obviously represents daily variability much better than long-term averages. However, the synthetic pan usually does not represent the full magnitude of particularly high or low observations. This is discussed in more detail in Section 9.2.

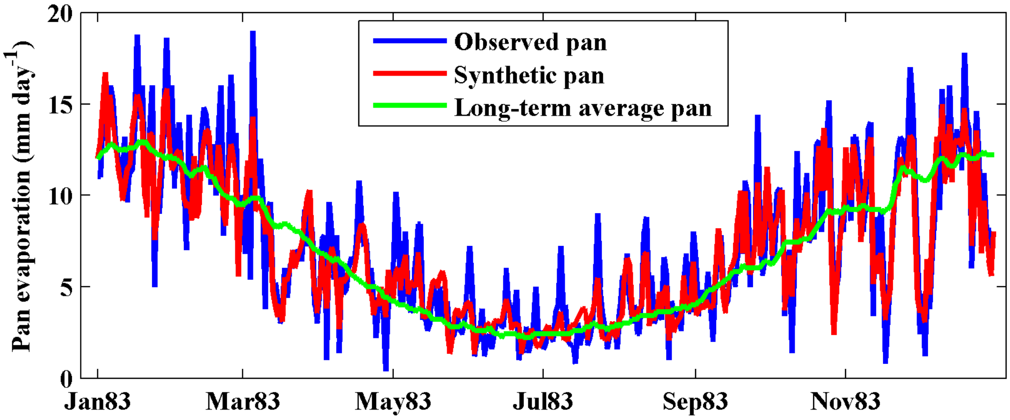

Figure 7.1: Kalgoorlie Post Office (12038) pan evaporation for 1983, a year with significant evaporation variability. The blue curve shows the observed pan evaporation, the red curve shows the synthetic pan evaporation, and the green curve shows the long-term average values for Kalgoorlie. The synthetic pan evaporation is a better representation of daily pan evaporation than the long-term averages, but underestimates the largest variations.

7.2 Improved overall accuracy.

The synthetic pan evaporation estimates observed pan more accurately than do the long-term averages. Comparing the mean-absolute-error for the synthetic pan shown in Figure 7.2 with the mean-absolute error for long-term average pan (Figure 7.3), both calculated using the validation dataset described in Section 6.1.1, demonstrates the improvement. For both datasets, the errors are highest in inland areas in December-February, where the evaporation rates are highest.

Note that the estimating accuracy using comparison against observed pan ignores any errors in observed pan evaporation record (see Sections 8.1 and 8.2).

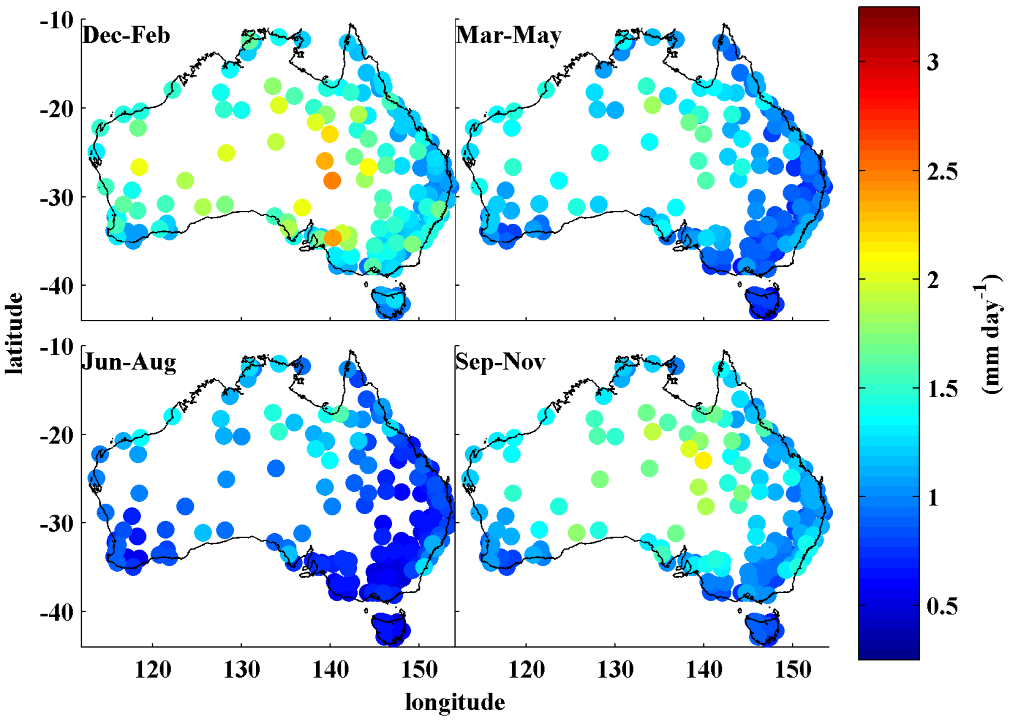

Figure 7.2: Mean-absolute-error using synthetic pan evaporation to estimate observed pan evaporation, for each season. Comparison with Figure 7.3 shows that the synthetic pan evaporation is a better estimate of observed pan evaporation than long-term averages. The errors are typically 30% lower than the long-term averages for December–February, 20% lower for March–May, 10% lower for June–August, and 25% lower for September–November.

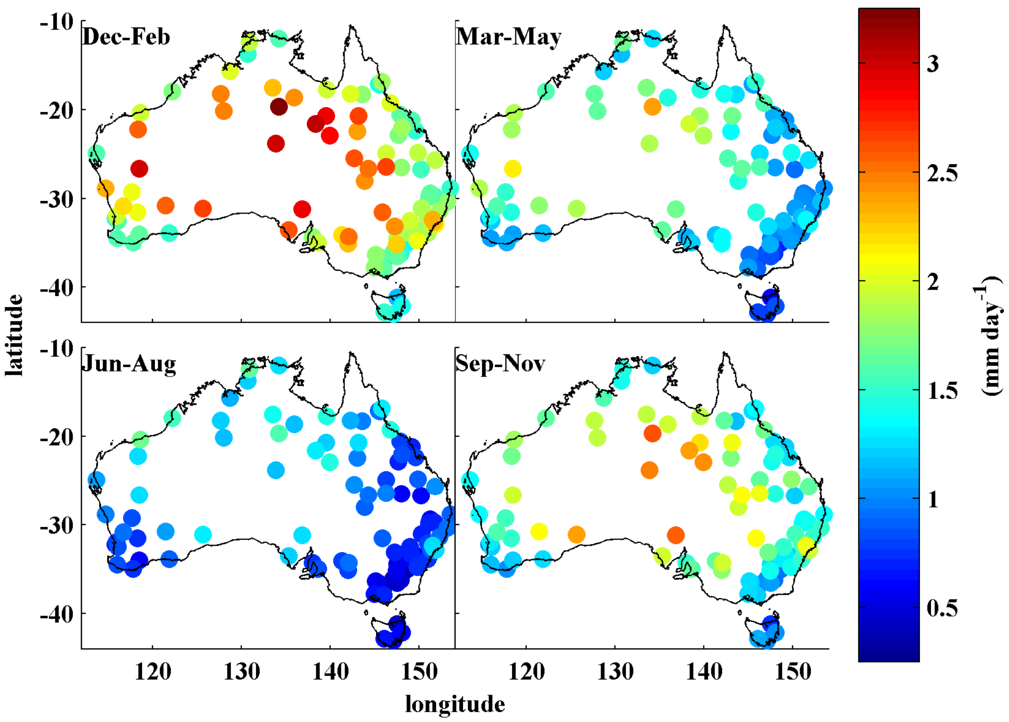

Figure 7.3: Mean-absolute-error using long-term averages to estimate observed pan evaporation, for comparison with Figure 7.2.

7.3 Improved relationship with rainfall.

The relationship between pan evaporation and rainfall is important in many modelling applications. For example, in hydrology, modelled runoff and drainage are often higher if daily evaporation data are used than if long-term average data are used. This is because drainage is typically modelled to occur after initial rainfall has saturated a soil moisture store. In the days following initial rainfall, the potential evaporation is likely to be lower than the long-term average, because of higher humidity, and possibly more clouds. So, when models are run using long-term average data, the post-rainfall evaporation from the soil moisture store is too high, and this must be replaced by follow-up rainfall before runoff or drainage occur. This can potentially lead to large systematic errors in water balance calculations[5].

A time-series analysis of rainfall and evaporation in the Border Rivers catchment, in the Murray-Darling catchment, demonstrated that the synthetic pan model represents the decrease in pan evaporation after significant rainfall quite well. The methodology follows Rayner et al. (2003). Firstly, catchment-average daily rainfall, observed[6] and synthetic pan evaporation were generated for the period 1975–2003. The pan evaporation time-series was expressed as anomalies from the monthly mean, to minimise seasonal effects. Each day in the period was categorised according to the number of days that had passed since a day of 15 mm catchment-average rainfall. The means of the observed and synthetic pan evaporation anomalies were then calculated for each rainfall category.

The results are shown in Figure 7.4. The average pan evaporation in the Border Rivers catchment is typically 37% below average when rainfall is 15 mm or higher, increasing back to the long-term average value over the following two weeks. Rainfall has a very similar effect in the observed and synthetic pan evaporation time-series. Obviously, rainfall has no effect on the long-term average time-series.

---

[5] Note that model calibration may mitigate the effects of such errors.

[6] The catchment-average "observed" rainfall and evaporation were calculated using the NR&M gridded surfaces.

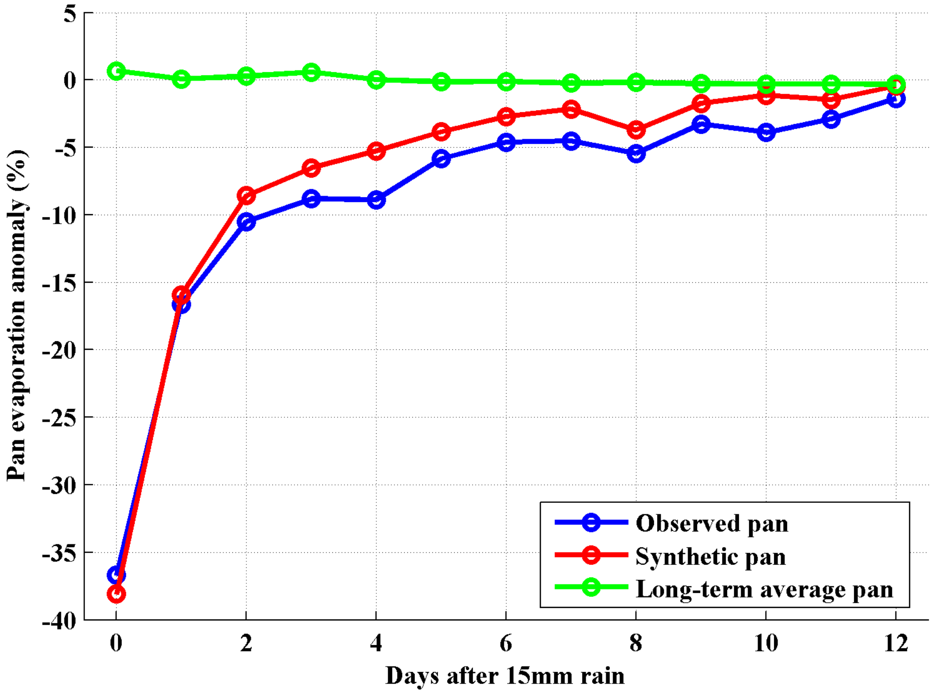

Figure 7.4: Rainfall-categorised evaporation anomalies for the Border Rivers catchment. Anomalies are departures from the average daily evaporation, calculated monthly. This figure demonstrates that the effect of rainfall on pan evaporation is captured by the synthetic pan, at least for this catchment. The small deviations in the long-term average pan anomalies are caused by the intra-month variations in the daily long-term average pan evaporation time-series.

8 Observed pan evaporation properties.

This section discusses some errors in the observed pan evaporation data.

8.1 Large errors.

A comparison of observed and synthetic pan evaporation for 2004 identified a number of records that were significantly different from the synthetic pan. Most of these discrepancies were associated with rainfall.

The observed pan evaporation records for all days in 2004 for which there was a difference of 15 mm or more between observed and synthetic pan evaporation are plotted in Figure 8.1. For most days, the station with the large difference is immediately obvious, because the observed pan evaporation is significantly higher than for neighbouring observations. These highly anomalous observations suggest that there are significant errors in the observed pan evaporation records (at least for recent records), even though the observations were flagged as "quality controlled and acceptable".

Of the 21 observations with evaporation differences of 15 mm or more, 15 observations occurred on days with 15 mm or higher rainfall, with 8 occurring on days with 50 mm or higher rainfall. On days when it rains, the rainfall (as indicated by a co-located rain gauge) must be removed from the pan before the evaporation is measured. Thus, any rainfall measurement error is transferred to the pan evaporation measurement. On days when the rainfall is much higher than the pan evaporation, a comparatively small error in the measured rainfall becomes a large error in the measured evaporation. This is probably the reason for these anomalous pan evaporation records.

It is also possible that rain may overflow the evaporation pan. Although evaporation observations made in such circumstances should be flagged as unacceptable, it is conceivable that some such records will slip through the various quality assurance processes.

The most pragmatic solution is probably to identify and remove the anomalous evaporation records before they are used to generate the NR&M gridded evaporation surfaces. However, this may produce errors in water-balance applications, if rainfall measurement errors are largely cancelled by transferring the error to the pan evaporation measurement.

The issue will certainly be investigated further.

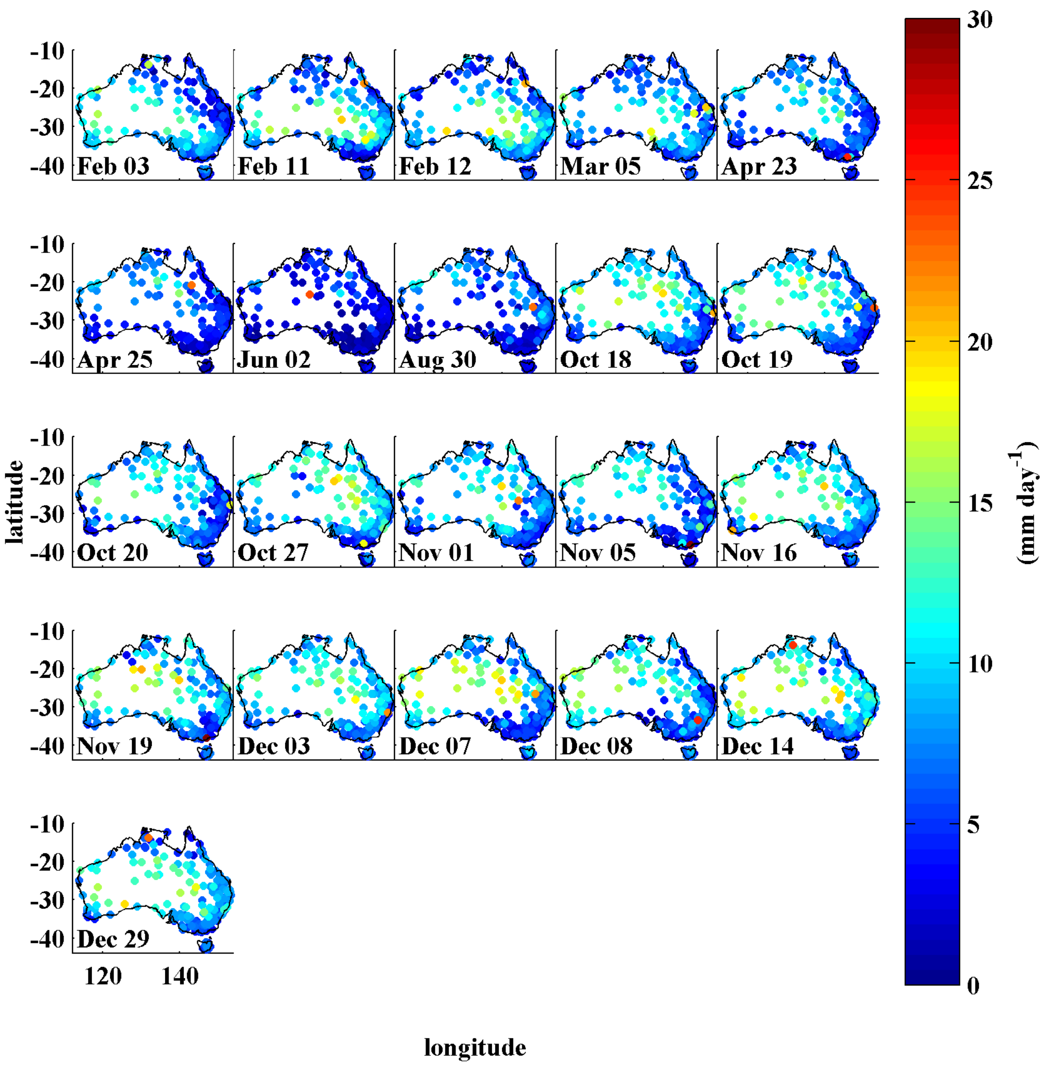

Figure 8.1: The observed pan evaporation for every day in 2004 for which there was a difference of 15 mm or more between an observed and synthetic pan evaporation record. For most days, the station with the large difference is immediately obvious, because the observed pan evaporation is significantly higher than for neighbouring observations. Most of these discrepancies are associated with rainfall.

8.2 Rounding of pan evaporation observations.

The observed Class A pan evaporation data supplied by ABM show the clustering at 4.0, 8.0, 12.0, 16.0 and 20.0 mm identified by Robinson (1998). This is related to observers measuring evaporation in “whole units” of the measuring tube, which holds 4 mm. The distribution of daily pan evaporation records is shown in Figure 8.2. The “dip” in the distribution above the 4 mm and 8 mm readings, but not below, suggests that observations may be preferentially rounded down, rather than simply rounded to the nearest 4 mm increment.

Rounding the observations to the nearest 4 mm should not significantly affect the calibration of the synthetic pan model. Further, the synthetic pan data do not show the clustering at 4 mm increments, even for stations where the observed data are highly clustered. I have not investigated whether synthetic pan data should be preferred to observed data for such stations.

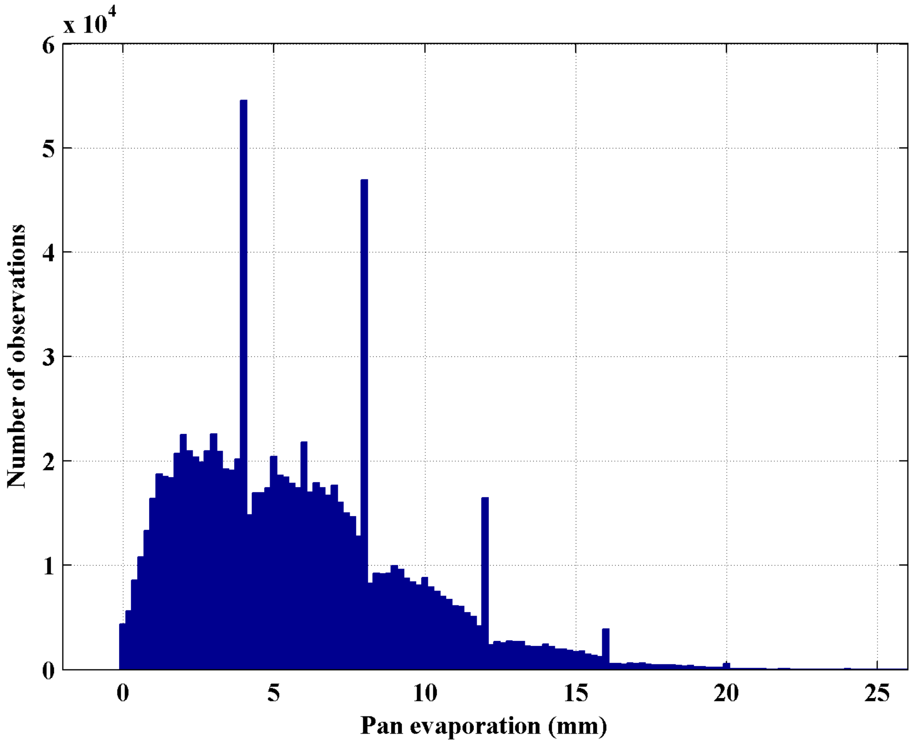

Figure 8.2: Distribution of daily Class A pan evaporation records, 1975–2003, showing clustering at 4.0, 8.0, 12.0, 16.0 and 20.0 mm.

9 Synthetic pan limitations.

The analyses in Section 7 demonstrated that synthetic pan is an improvement on long-term average pan. In this section, synthetic pan is compared to observed pan in more detail, and the limitations of the model discussed.

9.1 The synthetic pan model does not include wind.

Wind increases pan evaporation on dry days by moving water vapour away from the surface of the pan. The synthetic pan model does not include wind, because there is no reliable, readily-available, gridded daily wind dataset for Australia. However, observed wind run[7] data are available from the ABM on the CD-ROM "Australian Daily Wind Data IDCJDC06", and 113 stations with long, stable wind run records and matching pan evaporation records were selected from this dataset for analysis.

Figure 9.1 demonstrates that the synthetic pan error is highest on days with high wind run and/or high vapour pressure deficit, indicating that wind run would be a valuable additional input to the synthetic pan model. The less dramatic increase in the error for days with high vapour pressure deficit and *low* wind run is probably because, with wind omitted, the *c* coefficient derived in the regression model is implicitly derived for average wind conditions. The average wind speed is 170 km day^{-1} for this dataset, consistent with the wind speed at which the mean-absolute-error in Figure 9.1 is lowest.

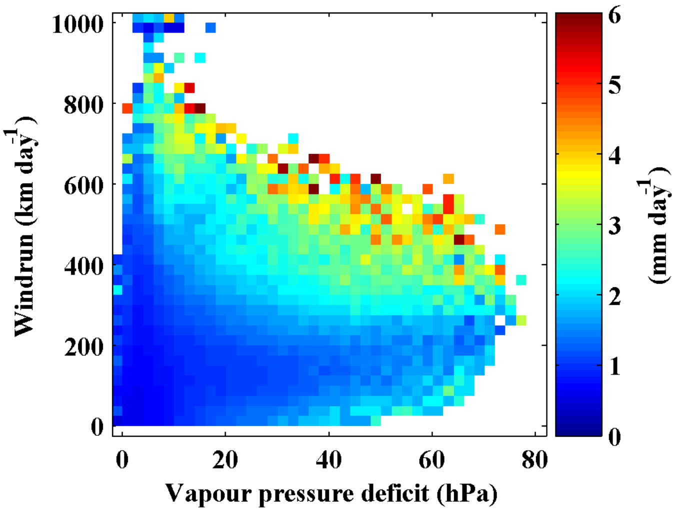

Figure 9.1: The mean-absolute-error in synthetic pan as a function of wind run and vapour pressure deficit. The error is highest on days with high wind run and/or high vapour pressure deficit, indicating that windrun would be a valuable additional input to the synthetic pan model.

9.2 The synthetic pan model does not represent high evaporation rates.

The poor representation of the highest and lowest daily pan evaporation observations was introduced in Section 7.1. A data-density plot showing the relationship between synthetic and observed pan evaporation in more detail is shown in Figure 9.2. Note the density scale has been normalised to emphasise high evaporation values. This figure shows that synthetic pan slightly

---

[7] "wind run" is average daily wind speed (km h^{-1}) × 24.

over-estimates low observed pan (by ~1 mm day^{-1}), but significantly underestimates observed pan evaporation records > 15 mm day^{-1}.

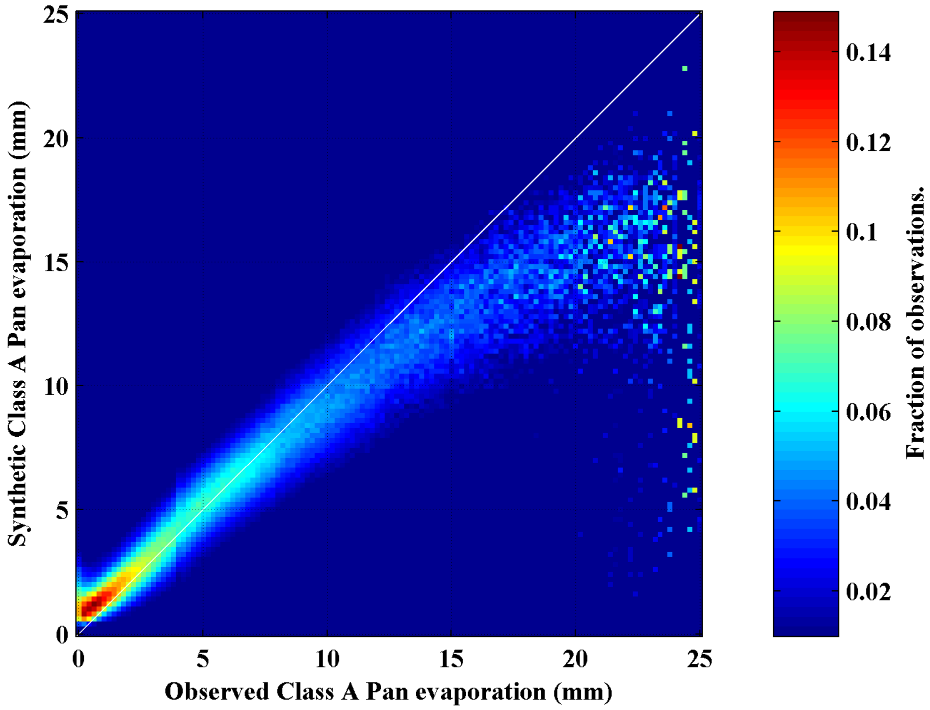

Figure 9.2: The relationship between observed and synthetic pan evaporation. The synthetic pan slightly over-estimates low observed pan (by ~1mm), but significantly underestimates high observed pan values. The density scale has been normalised to unity for each observed evaporation bin to highlight the high observed pan values; in reality, 99% of the pan evaporation records in the validation dataset are < 16 mm day^{-1}.

Several factors probably cause the under-estimation of high observed pan evaporation values, and the over-estimation of low values.

- The synthetic pan does not include wind; the days with the highest pan evaporation are presumably days with high solar radiation, high vapour pressure-deficit and high wind.
- The synthetic pan is derived from the NR&M gridded data surfaces, which were calculated using a thin-plate spline. Underestimating high values, and overestimating low values, is a known issue with the smoothing spline algorithm.
- The NR&M gridded solar radiation surfaces are derived from observations of 9 am and 3 pm cloud amounts (Jeffrey et al. 2001). This algorithm is known to systematically under-estimate high solar radiation values and over-estimate low values. However, this error should be mitigated by the regression algorithm used to derive the synthetic pan model.
- Rainfall is incorporated into the synthetic pan model implicitly through its relationship with solar radiation and vapour pressure deficit (Section 7.3). However, it rained on many of the days with low observed pan evaporation in Figure 7.1, which suggests that the effects of rainfall are not completely represented in the synthetic pan model.

9.3 Synthetic pan does not represent trends in observed pan evaporation.

Roderick and Farquhar (2004) found statistically significant trends in pan evaporation over the last 30 years for many Australian observing stations. The attribution of these trends to changes in more fundamental properties of the Australian climate has generated considerable debate within the Australian research community.

If trends in the synthetic pan replicated trends in the observed pan evaporation, then synthetic pan would be a powerful tool for attribution studies. Figure 9.3 compares trends in observed and synthetic pan evaporation. Unfortunately, trends in the observed pan evaporation are not represented by the synthetic pan evaporation. Possible reasons for this are:

- The trends may be caused by changes in atmospheric properties to which the synthetic pan is not sensitive, such as wind or changes in atmospheric opacity unrelated to cloud oktas.
- The trends may be caused by changes in very high or low evaporation rates, which are not represented well by the synthetic pan model.
- At least some of the trends may be due to changes in the local environment around the pan. Because the synthetic pan is calculated using interpolated surfaces, it should be less sensitive to local changes.

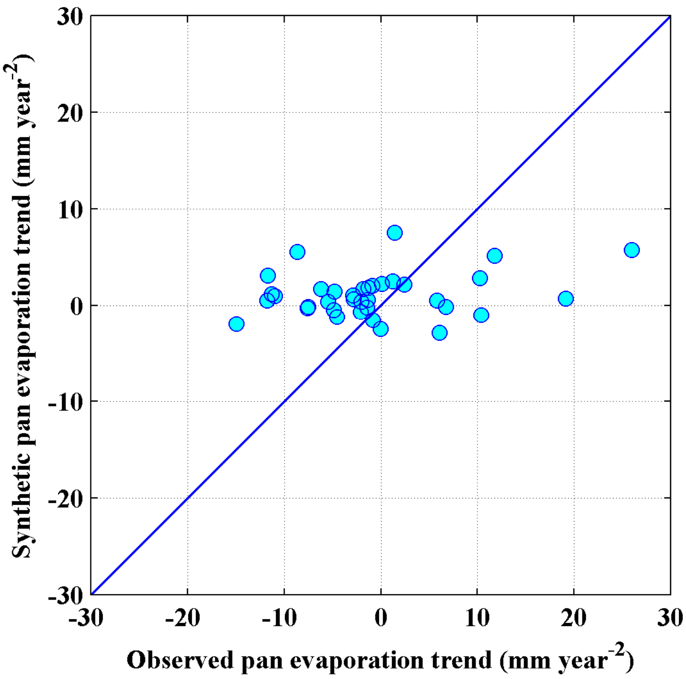

Figure 9.3: Comparison of trends in observed and synthetic pan evaporation for 1975–2003. Each point represents the trend for a single pan observing station. The trends in the observed pan evaporation are not represented by the synthetic pan evaporation (the points do not lie on the blue line). Data are only plotted for stations with highly-complete observation records (for at least 80% of years there are observations for at least 90% of days in every month.).

9.4 The synthetic pan model does not change with time.

Applying model coefficients derived using 1975-2003 observations to create a pre-1975 synthetic pan record implicitly assumes that the relationship between pan evaporation, vapour pressure deficit and solar radiation (or, more strictly, cloud amount) does not change with time. This *is* a serious concern, because the synthetic pan model is an empirical model, not a physical model.

The 2004 pan evaporation observations were not used to derive the model coefficients, and so provide an independent validation dataset. Figure 9.4 shows the average error in the synthetic pan data compared with 2004 observation data. For most stations, the average error is acceptable (< 1 mm day^{-1}) although over-all there is a tendency for the synthetic pan to *overestimate* observed pan. The persistence of these biases through each month (not shown) suggests that the biases are not due to particular weather conditions, but represent small changes in the relationship between observed pan evaporation, vapour pressure deficit, and cloud in the climate surfaces. Such changes *are* expected, because the stations and climate data used to generate the climate surfaces change through time. eg. meteorological stations open and close, observations for particular climate variables commence or cease, and there are subtle changes in the exposure of instruments caused by vegetation changes or nearby development.

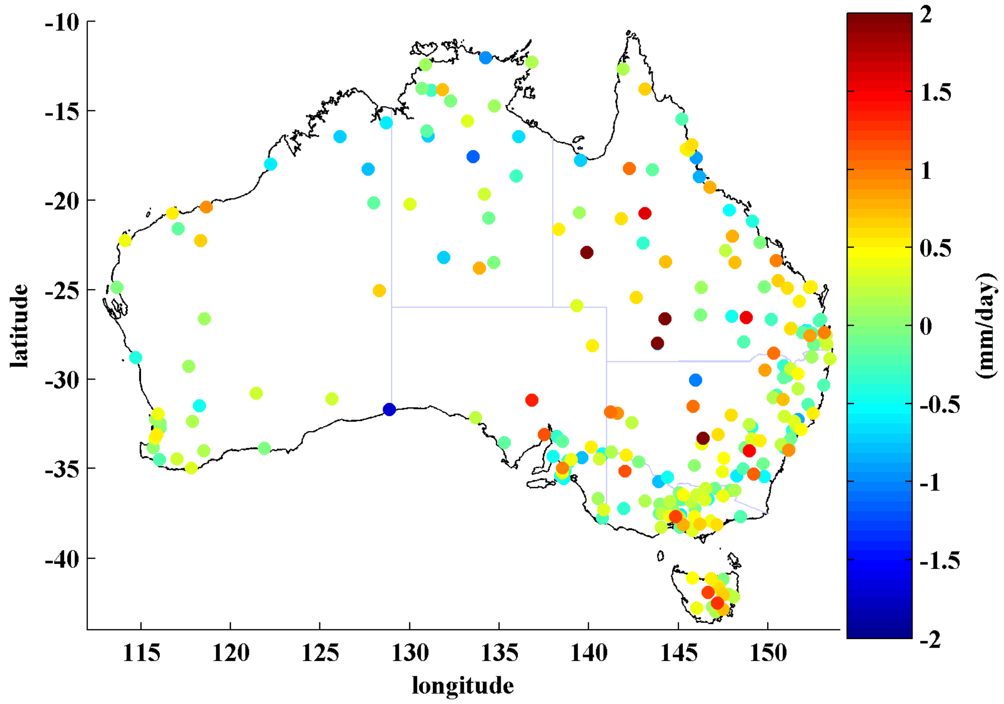

Figure 9.4: Average error in the synthetic pan evaporation data for 2004. The 2004 pan evaporation observations were not used to derive the model coefficients, and so provide an independent validation dataset. For the blue (positive) stations the synthetic pan overestimates the observed pan evaporation. Overall, there is a tendency for the synthetic pan to overestimate observed pan for 2004.

However, Figure 9.4 shows suspiciously large biases for a number of stations. Annual climate data for two stations with large discrepancies are shown in Figure 9.5 and Figure 9.6. In both cases, there is a large overall trend in the annual pan evaporation which is not consistent with the solar radiation and vapour pressure deficit time-series, nor with the rainfall time-series.

Station metadata supplied by ABM suggest some of these changes are related to the environment around the pan. The observing station at Boulia (38003, Figure 9.5) was relocated from a cluttered urban site at Boulia Post Office to an open site at Boulia Airport in July 1999, which is consistent with an increase in pan evaporation at this time. The site diagrams for Maningrida (14400, Figure 9.6) suggest that it has been relocated from an open area to an urban area sometime between 1973 and 1992. In addition, the pan has been replaced an unusually-high six times since its installation in 1966[8]. Thus, although no definitive cause for the lower evaporation

---

[8] In 1972, 1978, 1982, 1984 and 1996.

at Maningrida from 1996 to 2001 could be identified, given the urban environment and instrumental changes, the pan evaporation variations are not surprising.

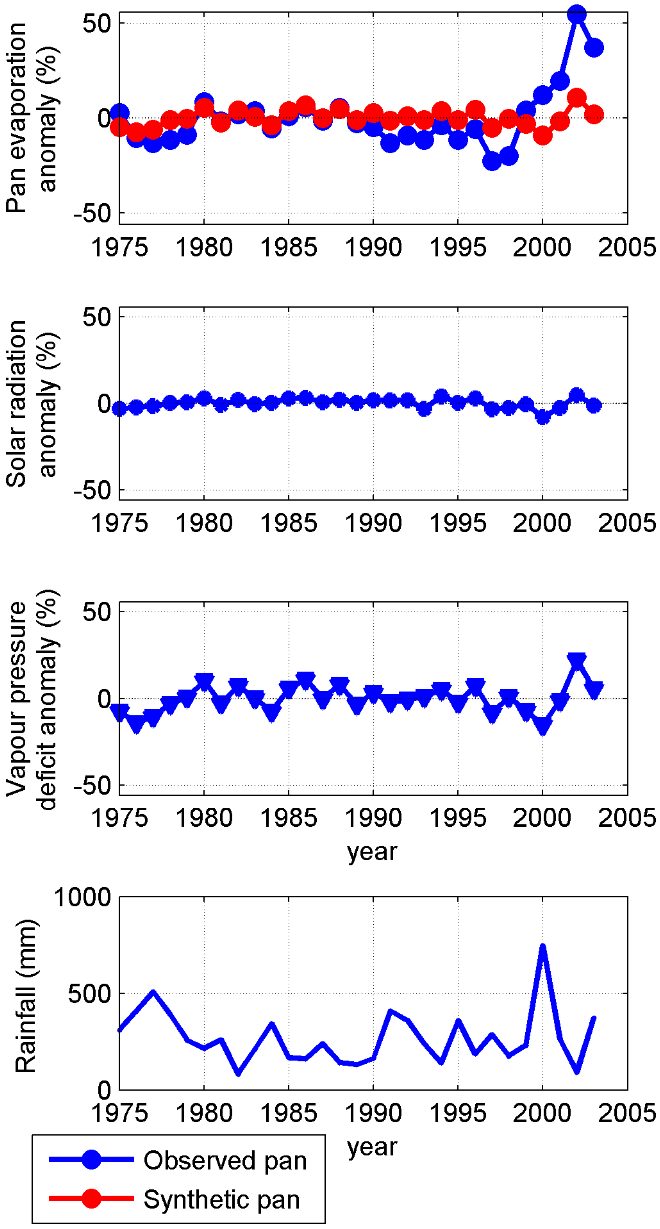

Figure 9.5: Annual time-series Boulia Airport, Qld (38003). The synthetic pan for this station has the largest negative error in Figure 9.4. Pan evaporation since 1999 has been significantly above average. The observing station at Boulia was relocated from a cluttered urban site at Boulia Post Office to an open site at Boulia Airport in July 1999, which is consistent with an increase in pan evaporation at this time. However, the anomalous decrease in observed pan evaporation between 1996 and 1997 remains unexplained.

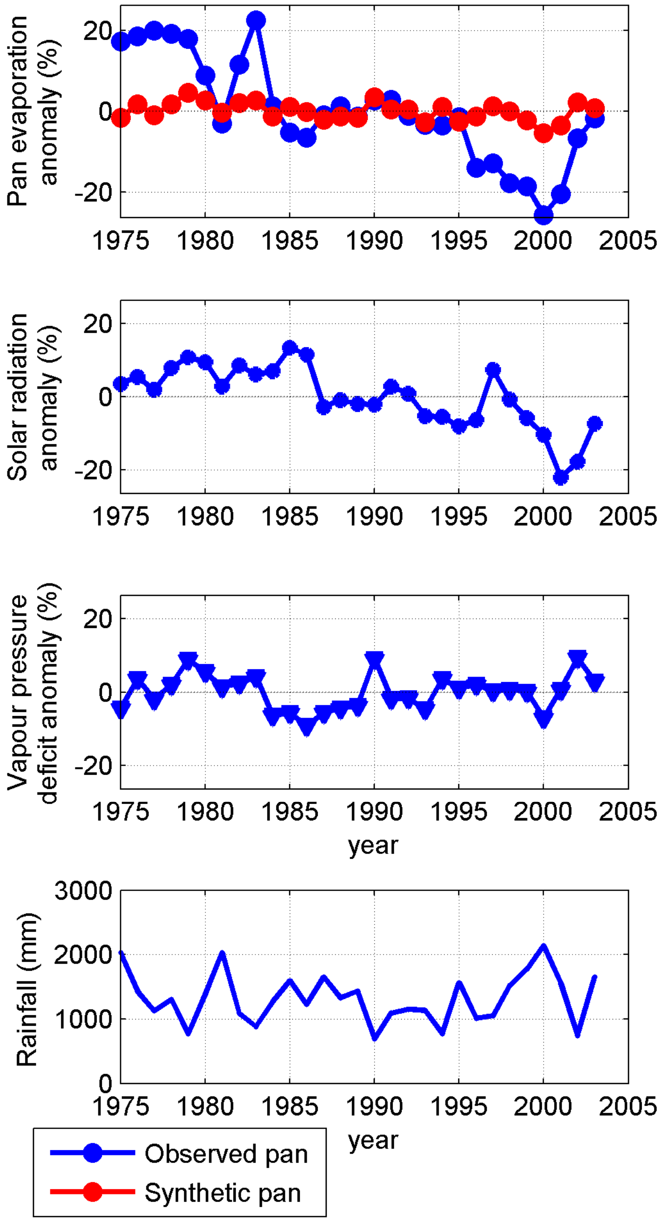

Figure 9.6: Annual climate time-series for Maningrida, NT (14400). The synthetic pan for this station has a large positive error in Figure 9.4. Site diagrams for the Maningrida observing station (not shown) suggest that it has been relocated from an open area to an urban area sometime between 1973 and 1992, which may explain the decrease in pan evaporation around 1980. No definitive cause for the lower evaporation from 1996 to 2001 could be identified.

9.5 Limitations specific to pre-1957 synthetic pan data.

The NR&M daily maximum and minimum temperature, solar radiation and vapour pressure surfaces for the period 1889-1956 were computed using an anomaly-interpolation spline algorithm (Rayner et al. 2004). The anomaly-interpolation spline method involved interpolating observed deviations from long-term average surfaces, rather than interpolating climate data directly. An important aspect of the method was to insert zero-anomalies into data-sparse regions prior to interpolation. The use of zero-anomalies prevented unrealistic values being generated for data-sparse regions, even when the interpolation was run with very few input stations.

As a consequence of using zero-anomalies, the *observation signal* in the climate surfaces – the extent to which the surfaces represent actual daily climate observations rather than long-term averages – is not spatially or temporally uniform. Instead, the surfaces gradually revert to long-term averages as the number of observations decreases in the early period of the climate record.

The gradual decrease in the observation signal of the input climate surfaces means that the day-to-day variations in early synthetic evaporation time-series will also be lower. The plots in Appendix C give an indication of where and when the synthetic evaporation surfaces represent day-to-day climate variability. The plot for each year shows the ratio of the root-mean-square (*rms*) daily variation, for that year, to the median *rms* daily variation for 1957 to 2003. Both the seasonal and inter-annual variations were subtracted prior to the analysis.

The *dark regions* are where there is little daily variation, and hence the surfaces represent long-term average climatic conditions. The *light regions* are where the daily variation is similar to the 1957 to 2003 surfaces. *Regions which are white* have at least 80% of the median daily variation in the 1957 to 2003 surfaces.

Finally, note that the images only indicate the amount of day-to-day variability in the synthetic evaporation surfaces. They should not be taken as an indication of data-quality.

10 Future improvements.

Solar radiation derived from cloud amounts and simple daily-average vapour pressure deficit together account for more than 80% of the total temporal and spatial variability in observed daily pan evaporation across Australia. This will probably be adequate for many biophysical modelling applications. However, the synthetic pan model will not be adequate for applications which need an accurate representation of extreme evaporation rates, or for analysis of pan evaporation trends.

To create a pan evaporation model suitable for climate change analysis, the following factors need to be considered:

- A high-quality observed pan evaporation dataset is required. Creating a universally-applicable pan evaporation model which correctly represents observed trends will be impossible without detailed, historical information about the pan environment and observing practices. Such information will either have to be parameterised and included in the synthetic pan model, or else a smaller "homogeneous" dataset, consisting only of stations without complicating local factors, will need to be complied.
- Reliable, gridded, daily, historical wind data are required. As demonstrated in Section 9.1, the error in the synthetic pan is high on days with high wind-run. In coastal regions, the direction of the wind, and the variation in wind throughout the day, should also be investigated.
- The effect of rainfall on pan evaporation should be examined in greater detail.
- Changes in the heat energy of the pan water could be included in the model.
- Daily pan evaporation is the integration of diurnal variations in air temperature, solar radiation, wind, humidity and pan water temperature. There must be a limit to how well this integral can be modelling using daily climate data. It would be very useful to have some idea what this limit is.
- The model should include both solar radiation and long-wave radiation.
- Finally, all the above factors need to be synthesised into a physically-based model for pan evaporation, along the lines of Linacre (1994).

11 References.

Allen, R.G., Pereira, L.S., Raes, D. & Smith, M. 1998, *Crop evapotranspiration - Guidelines for computing crop water requirements*, 2nd edn, Food and Agriculture Organization of the United Nations, Rome.

Carter, J.O., Hall, W.B., Brook, K.D., McKeon, G.M., Day, K.A. & Paull, C.J. 2000, "Aussie GRASS: Australian grassland and rangeland assessment by spatial simulation" in *Applications of Seasonal Climate Forecasting in Agricultural and Natural Ecosystems*, eds. G.L. Hammer, N. Nicholls & C. Mitchell, Kluwer Academic Publishers, The Netherlands.

Hounam, C.E. 1961, "Evaporation in Australia: A critical Survey of the Network and Methods of Observation together with a Tabulation of the Results of Observations", *Bureau of Meteorology Bulletin*, vol. 44, pp. 1-32.

Jeffrey, S.J., Carter, J.O., Moodie, K.B. & Beswick, A.R. 2001, "Using spatial interpolation to construct a comprehensive archive of Australian climate data", *Environmental Modelling and Software*, vol. 16, no. 4, pp. 309-330.

Linacre, E.T. 1994, "Estimating US Class A pan evaporation from few climate data.", *Water International*, vol. 19, pp. 5-14.

Penman, H.L. 1948, "Natural evaporation from open water, bare soil and grass", *Proc.R.Soc.London, Ser.A*, vol. A193, pp. 120-145.

Rayner, D.P., Moodie, K.B., Beswick, A.R., Clarkson, N.M. & Hutchinson, R.L. 2004, *New Australian daily historical climate surfaces using CLIMARC.* Queensland Department of Natural Resources, Mines and Energy, Australia.

Richards, R., Watson, I., Bean, J., Maconochie, J., Clipperton, S., Beeston, G., Green, D. & Hacker, R. 2001, *Australian Grassland and Rangeland Assessment by Spatial Simulation (Aussie GRASS). Southern Pastures Sub-project*, Queensland Department of Natural Resources and Mines, Australia.

Robinson, J.B. 1998, "Bias in Daily Pan Evaporation Data", *Agricultural Science*, vol. 1, no. 4, pp. 22-24.

Roderick, M.L. & Farquhar, G.D. 2004, "Changes in Australian pan evaporation from 1970 to 2002", *International Journal of Climatology*, vol. 24, pp. 1077-1090.

van Dijk, M.H. 1985, "Reduction in evaporation due to the bird screen used in the Australian class A pan evaporation network", *Australian Meteorological Magazine*, vol. 33, pp. 1

12 Appendix A: Regression Coefficients.

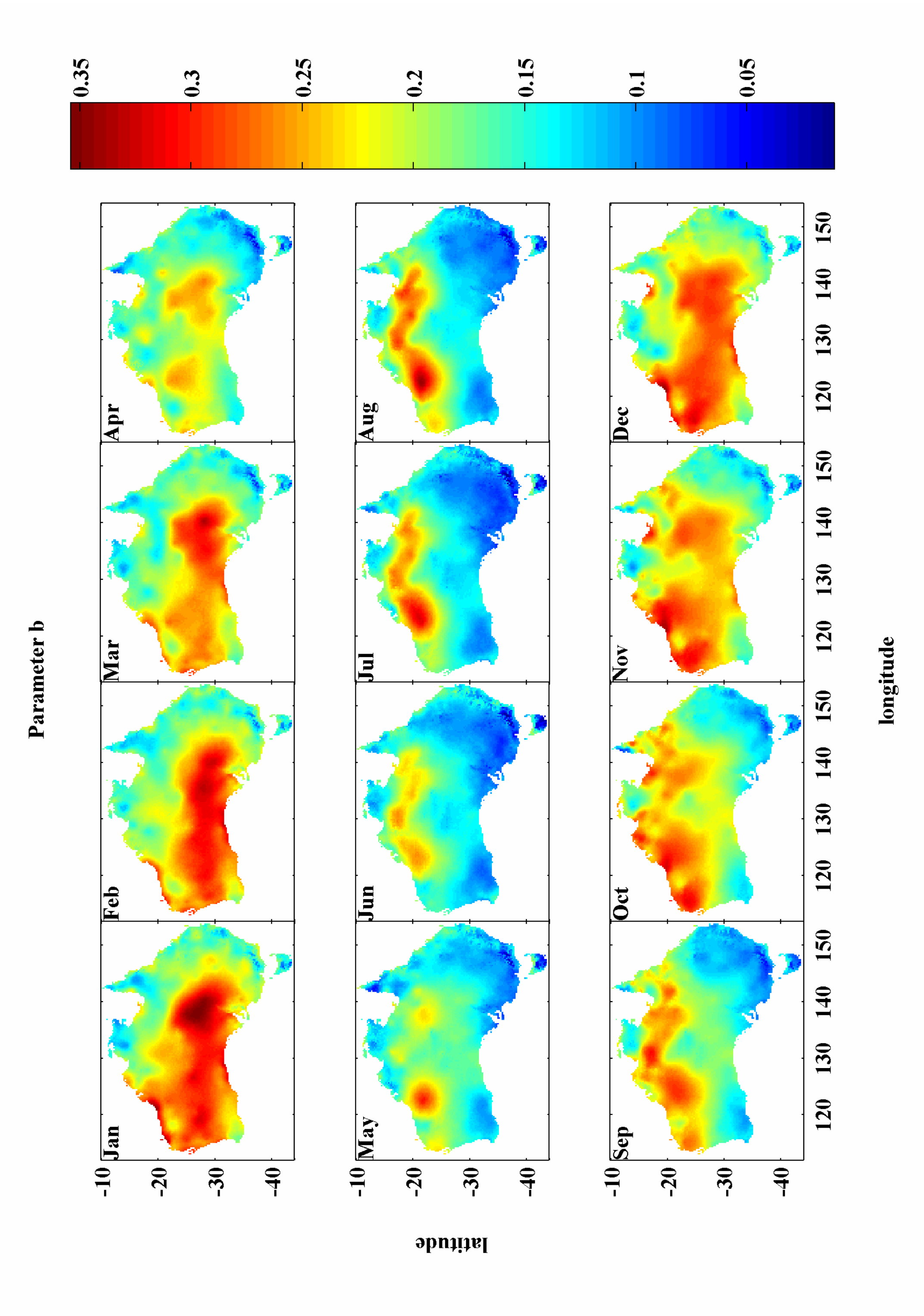

Figure 12.1: Synthetic pan evaporation regression parameter b. Recall $E_{pan} = b \times R_s + c \times (e_s - e_a)$.

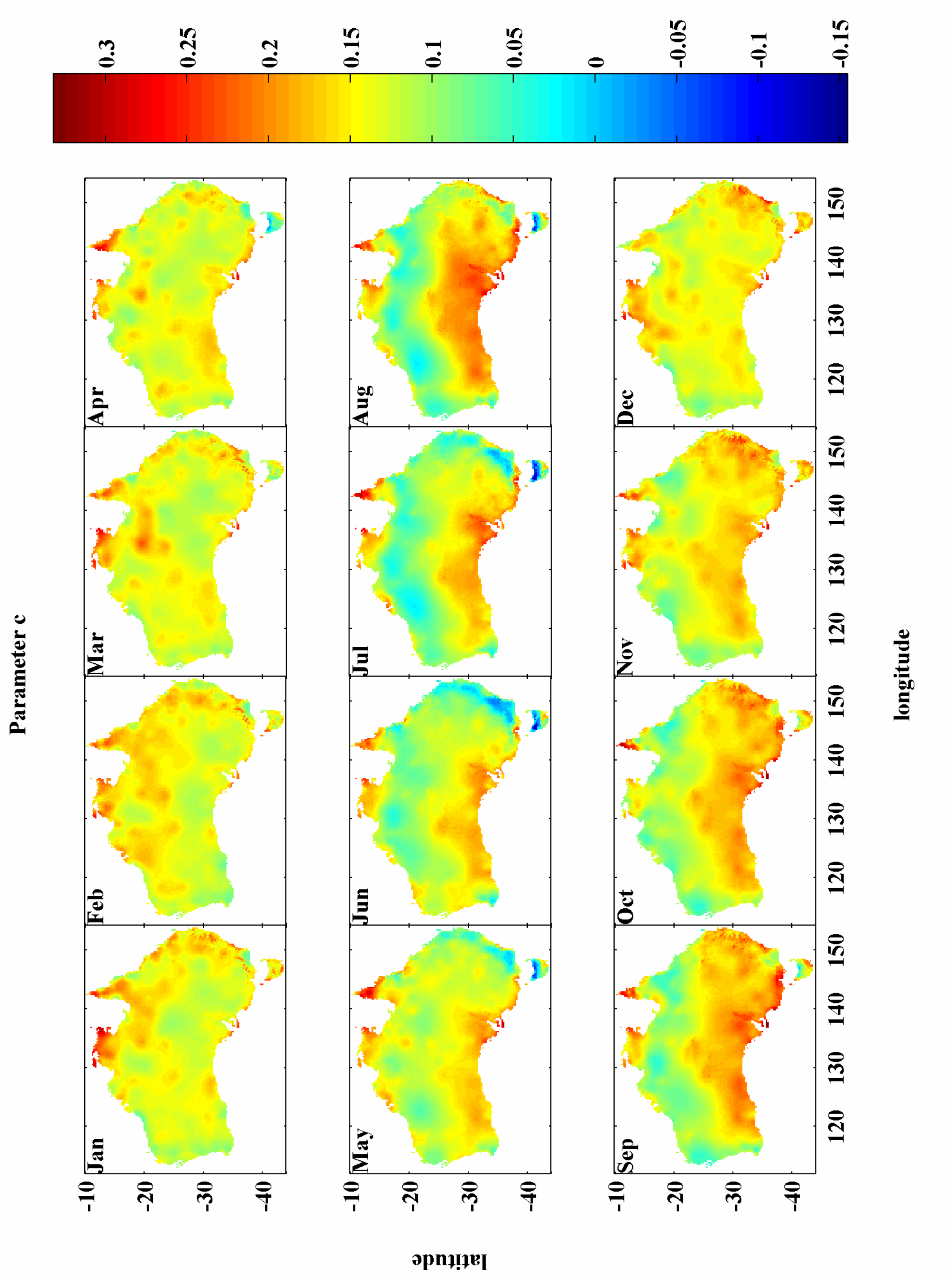

Figure 12.2: Synthetic pan evaporation regression parameter c. Recall
$E_{pan} = b \times R_s + c \times (e_s - e_a)$.

13 Appendix B: Synthetic evaporation model components.

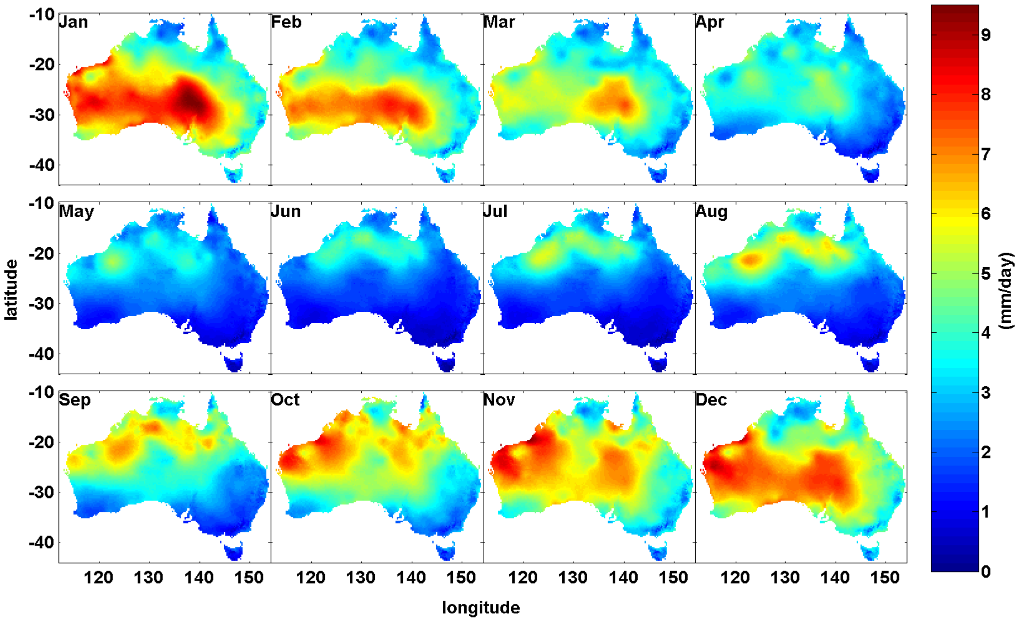

Figure 13.1: Mean solar radiation term $b \times R_s$ for each month, calculated using data for 1975-2003.

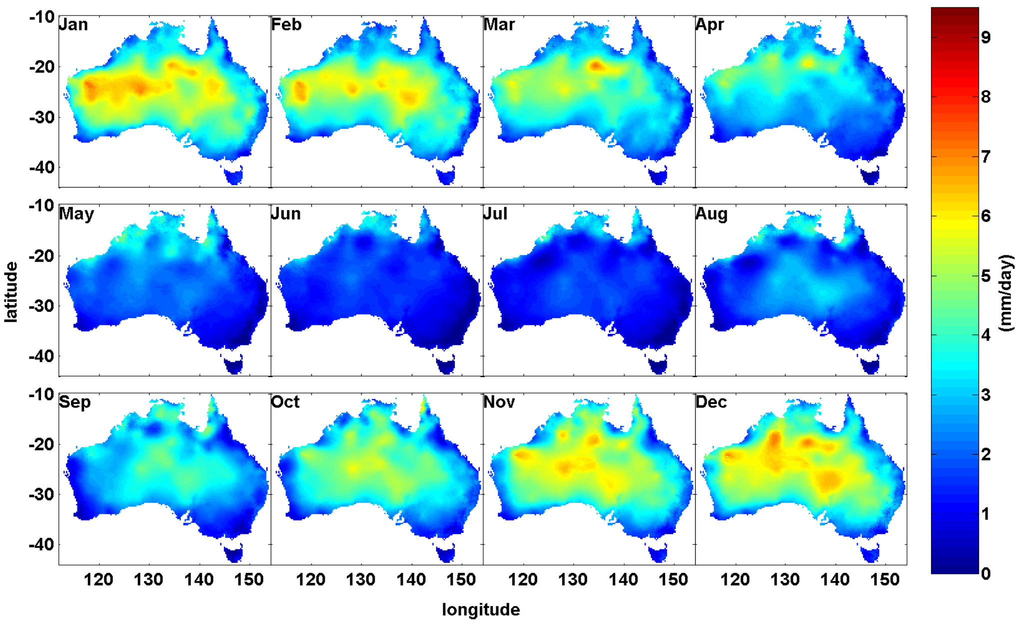

Figure 13.2: Mean vapour pressure deficit term $c \times (e_s - e_a)$ for each month, calculated using data for 1975-2003.

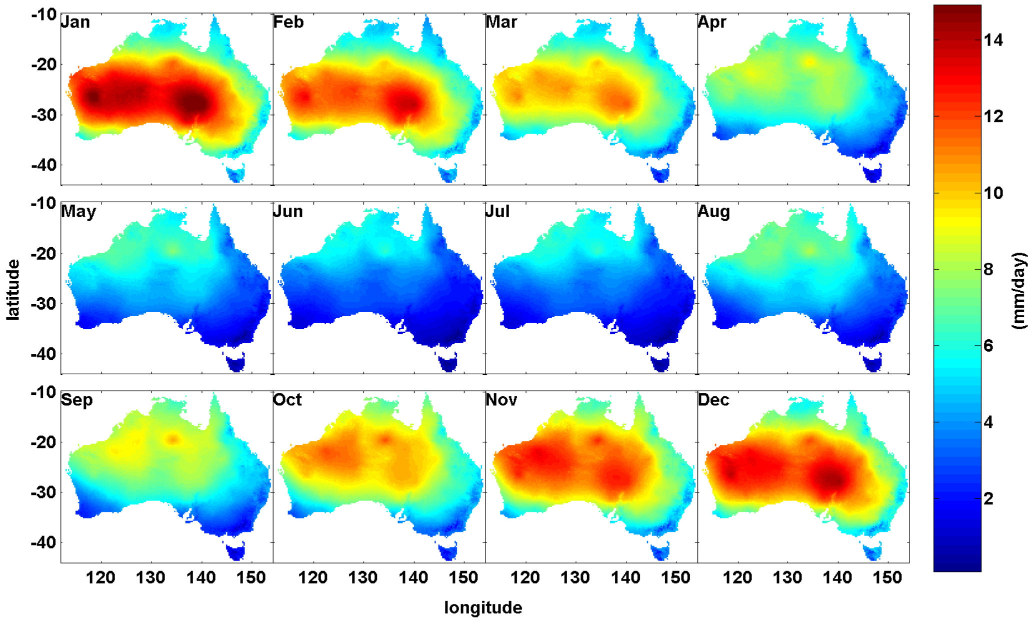

Figure 13.3: Mean synthetic pan evaporation $E_{pan} = b \times R_s + c \times (e_s - e_a)$ for each month, calculated using data for 1975-2003.

14 Appendix C: Day-to-day variations in the synthetic pan evaporation surfaces.

These plots give an indication of where and when the synthetic evaporation surfaces represent day-to-day climate variability. The plot for each year shows the ratio of the root-mean-square (*rms*) daily variation, for that year, to the median *rms* daily variation for 1957 to 2003. Both the seasonal and inter-annual variations were subtracted prior to the analysis.

The *dark regions* are where there is little daily variation, and hence the surfaces represent long-term average climatic conditions. The *light regions* are where the daily variation is similar to the 1957 to 2003 surfaces. *Regions which are white* have at least 80% of the median daily variation in the 1957 to 2003 surfaces.

Finally, note that the images only indicate the amount of day-to-day variability in the synthetic evaporation surfaces. They should not be taken as an indication of data-quality.

14.1 Synthetic pan evaporation.

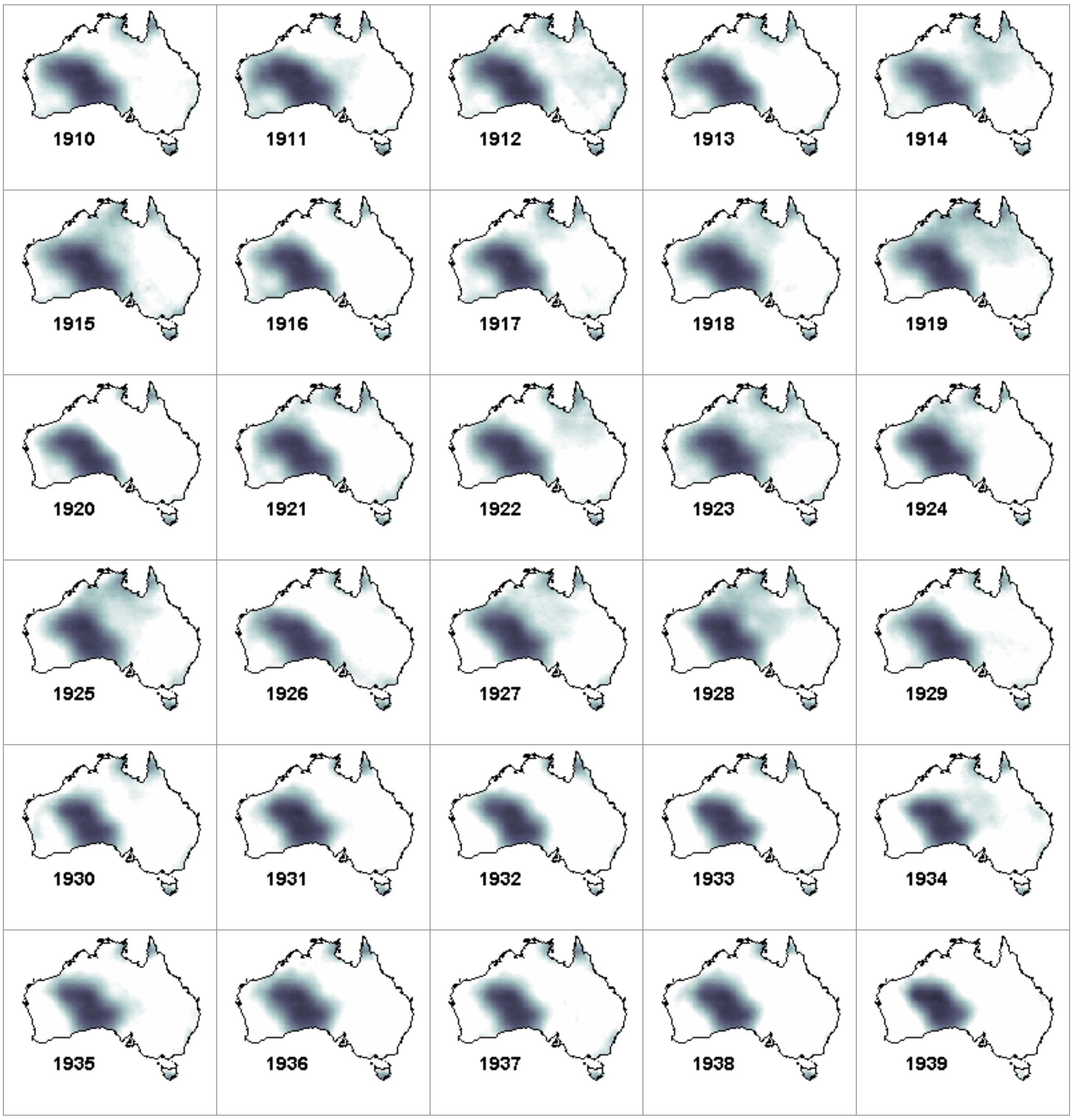

Figure 14.1: Day-to-day variation in the synthetic pan evaporation surfaces.

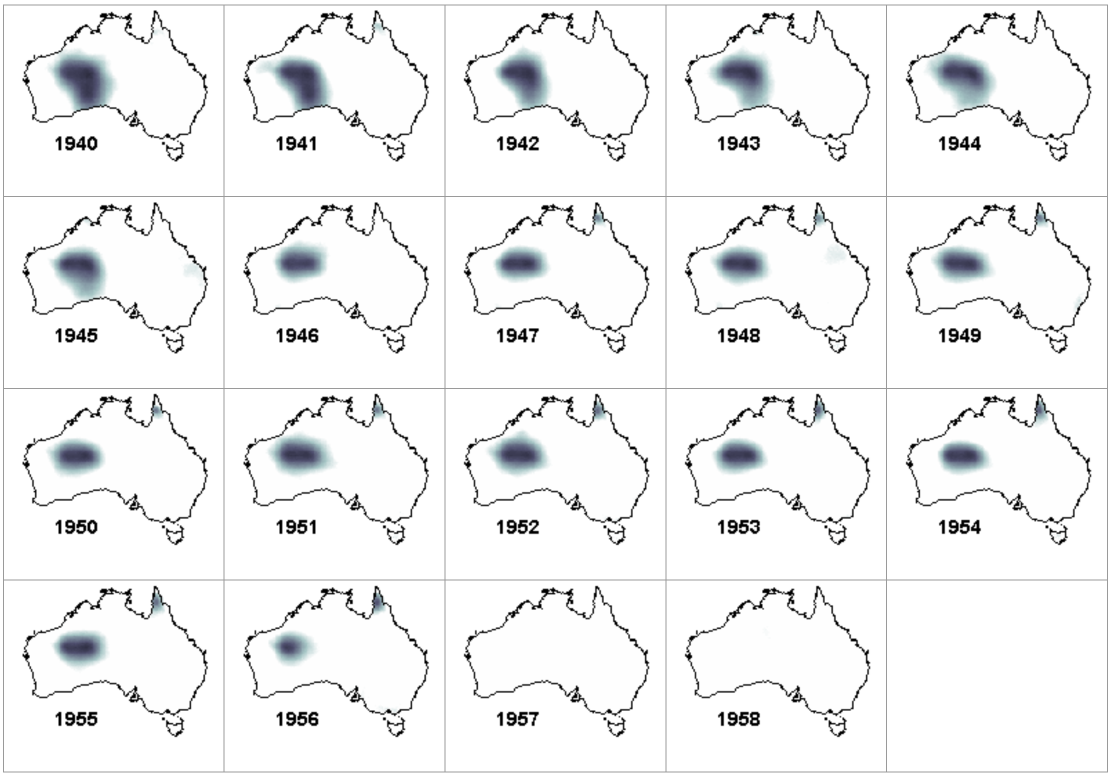

Figure 14.1: Day-to-day variation in the synthetic pan evaporation surfaces. (cont.)

15 Appendix D: The adequacy of the NR&M gridded solar radiation.

15.1.1 Summary.

The NR&M solar radiation grids are calculated from 9 am and 3 pm observations of cloud amounts (Jeffrey et al. 2001). A simple analysis using high-quality ground-based observations of solar radiation demonstrated that these grids are adequate for the current synthetic pan model.

15.1.2 Backgroud.

The NR&M solar radiation grids are calculated from 9 am and 3 pm observations of cloud amounts. There are daily cloud observation records for hundreds of stations, with data covering most of the 20th century (although most of the pre-1957 records are not available in electronic form). The combination of wide geographic coverage and long observation records which are relatively free from systematic changes makes cloud amounts preferable to ground-based pyronometer measurements, satellite observations or sunshine-duration records for estimating historical solar radiation.

A previous study compared the NR&M gridded solar radiation with data from the Australian Government Bureau of Meteorology's new surface network[9]. The NR&M gridded daily solar radiation was typically in error by 16% at the locations of the surface network sites, with low values being systematically over-estimated and high values being systematically under-estimated.

15.1.3 Comparison of synthetic evaporation with surface network solar radiation.

Despite the above study, I decided to check whether the error in the synthetic pan from using solar radiation modelled from cloud amounts was significant. The NR&M solar radiation grid residuals (grid value minus ABM surface network value) were compared with synthetic pan residuals (synthetic pan minus observed pan) for co-located evaporation stations, and the results are shown in Figure 15.1. For most stations, there is very little correlation, which indicates that improved solar radiation data would not significantly improve the current synthetic pan model. This contrasts with the correlation between wind run and synthetic pan residuals, shown in Figure 15.2. For some stations, wind run alone accounts for 50% of the synthetic pan residuals.

In summary, the errors in the synthetic pan evaporation caused by errors in the NR&M solar radiation grids are considerably less than the errors caused by not including wind information.

---

[9] On the CD-ROM "Australian Solar Radiation Data - NCCSOL 2.209"

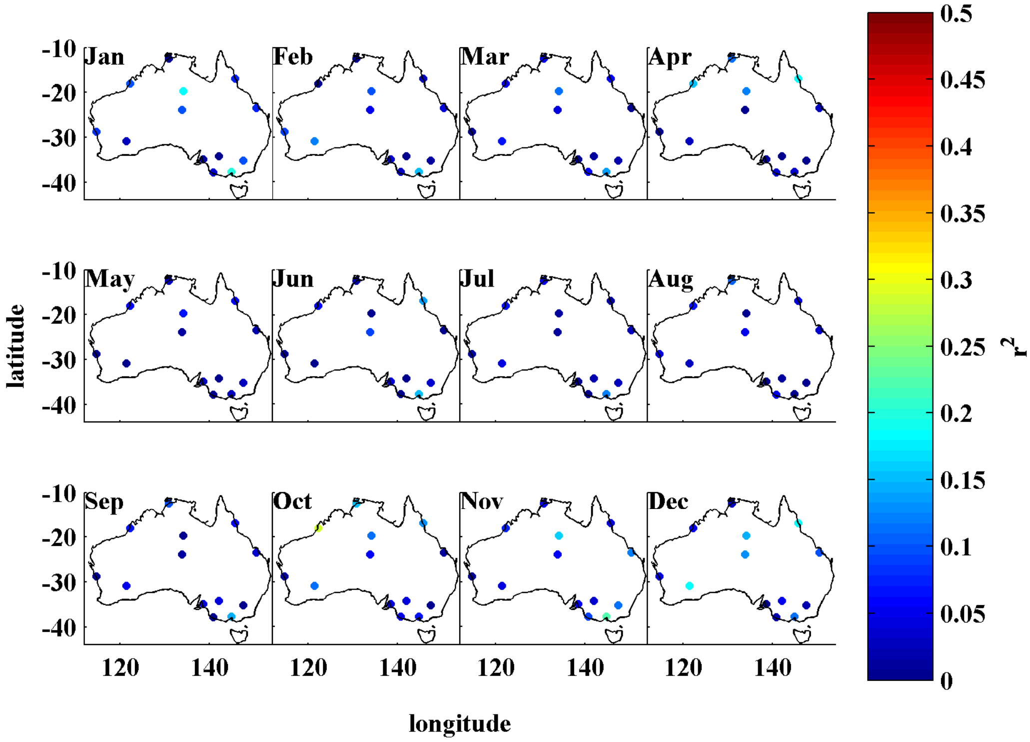

Figure 15.1: Correlation between synthetic pan residuals and NR&M solar radiation residuals. For most stations, there is very little correlation, which indicates that cloud-derived solar radiation is adequate for the current synthetic pan model.

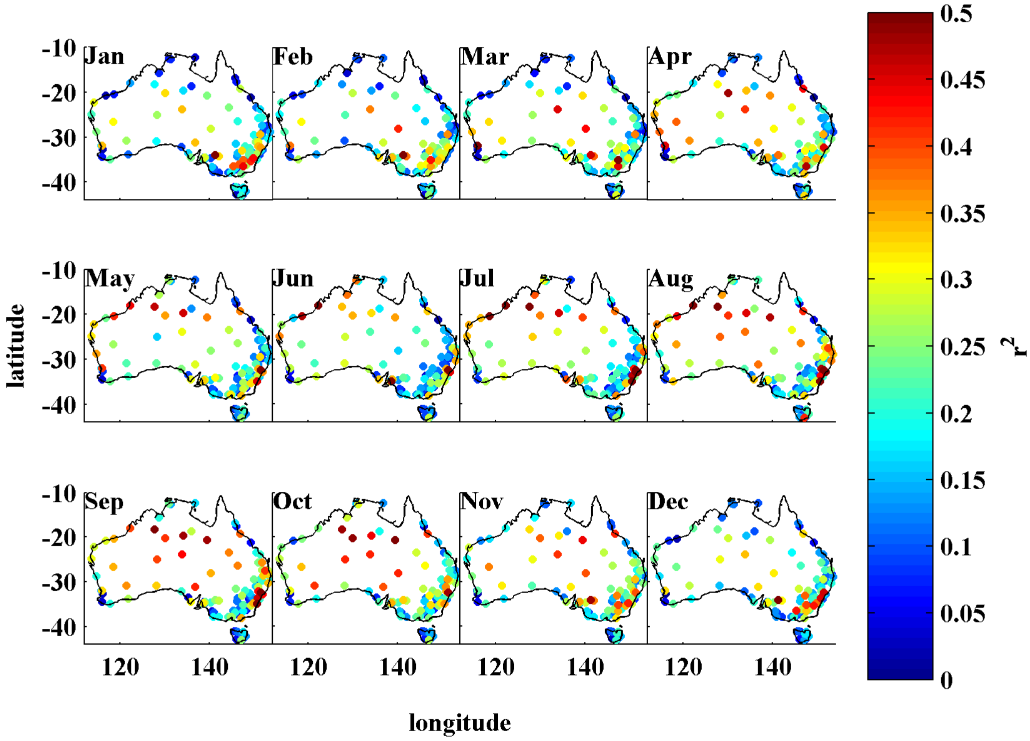

Figure 15.2: Correlation between synthetic pan residuals and wind run. The correlation is generally much higher than the correlation with errors in cloud-derived solar radiation shown in Figure 15.1.