Chapter 5

Evaluating Climate Model Simulations of Daily Rainfall Using Estimated Areal-Rainfall Parameters

Contents

5.1. Introduction

The mismatch in spatial scale between climate model and observed station time series of precipitation has, in the past, limited evaluation of climate models to more qualitative than quantitative comparisons.

The earlier chapters (3 and 4) of this thesis have presented the development of a technique for estimating the variability characteristics of an areal daily rainfall time series, using information from a limited number of available stations. Previous model evaluations have been limited by the incomparable levels of temporal variability between point observations and smoother areal model rainfall time series, or have had to rely on averages of available stations which may not represent a 'true' areal mean if an insufficient number or distribution of stations is available. The approaches applied here predict the parameters of a time series made up from an infinite number of stations in a box, using the characteristics of spatial and temporal variability in the stations that are available, and going some way to overcoming the scale mismatch.

In this chapter, the application of these novel techniques is demonstrated by the evaluation of GCM simulated precipitation for the UK and the northern region of South Africa. South Africa is chosen as a second case study region in order to demonstrate the application of this technique to a region where data coverage is sparser than for the UK. This second study region will also give some indication as to whether particular model errors are common to more than one region. The model evaluations shown in this chapter are illustrative of how the method can be applied, and not intended to be a thorough investigation into the dynamics of each model.

Firstly, some details of the models used are presented (**Section 5.2**), before the methods used are described (**Section 5.3**). **Section 5.4** presents the model evaluation results for the UK and **Section 5.5** the results for South Africa. The methods and results are then discussed and summarised in **Section 5.6**.

5.2. Model Details

The three models were selected from those made available through the Program for Climate Model Diagnosis and Inter-comparison (PCMDI)[1] for the IPCC AR4, which includes extensive archives of the newest generation coupled ocean-atmosphere general circulation models. Some details of each of the three models are given in Table 5-1.

Model Name	**Centre /Country**	**Atmospheric resolution**	**Flux correction**	**Cloud parameterisation**	**Convective Parameterisation**	**Key reference**
HadCM3	Met Office/ UK	2.5° x 3.75°	None	Large scale rainfall and snowfall calculations based on cloud water and ice contents (similar to Smith 1990)	Mass flux penetrative convection scheme (Gregory and Rowntree, 1990) with improvement by Gregory *et al* (1997)	Gordon *et al.* (2000)
CGCM3.1	CCCMA/ Canada	3.75° x 3.75°	Heat, water	Prognostic cloud evaluated through a relative humidity threshold; precipitation occurs whenever the local relative humidity is supersaturated.	Moist Convective adjustment scheme	Flato *et al.* (2000)
PCM.1	NCAR/ USA	2.8° x 2.8°	None	Prognostic cloud condensate with diagnostic cloud amount	Parameterization schemes developed by Zhang and McFarlane (1995) and Hack (1994)	Washington *et al.* (2000)

Table 5-1: Technical details of the 3 GCMs evaluated (Dai *et al.*, 2006, Sun *et al.*, 2006)

[1]Available at http://www-pcmdi.llnl.gov/#

5.3. Method

Grid boxes for the UK and South Africa, for each of the three GCMs studied, were evaluated for the period 1961-2000. For HadCM3, the full 240-year control run was taken to represent the climate of this period. The longer control simulation allows more robust estimates of statistics such as the 95^{Th} percentile value as a greater amount of data is available. For CGCM3 and PCM, such a control run was unavailable, and so the period 1961-2000 was taken from the 20^{th} century transient simulations. This means that less data is available to calculate the relevant statistics, but for CGCM3 an ensemble of three model runs was available, thus increasing the sample size threefold.

The following statistical parameters for the precipitation time series for these model simulations were compared to those which are estimated for the 'true' areal observed time series.

1) ***MD*** -Mean daily rainfall amount - arithmetic mean of all days for which values are recorded.
2) ***P(d)*** - Dry-day probability – number of days recording less than 0.3mm divided by the number of days recorded.
3) ***MWDA*** - Mean wet-day rainfall amount (Mean intensity) - mean daily rainfall divided by wet-day probability.
4) ***α* and *β*** - Gamma distribution shape and scale parameters fitted using an approximation to the maximum likelihood method.
5) ***P95*** - 95^{th} percentile value determined from the distribution fitted to wet days only. This is therefore the precipitation amount that is exceeded on 5% of wet-days, not on 5% of all days.

Simulated values of these parameters, for each season, are compared with estimated observed areal values. The methods by which each of these values is calculated are summarised as follows:

MD_N: mean of the mean daily rainfalls at all *n* available stations within the grid box.

$P(d)_N$: $P(d)_N = [\overline{P(d)_{i,N}}]^{N'}$ (where $N' = 1/\overline{r(w/d)}$, when $N=\infty$)

$MWDA_N$: $MD_N/(1-P(d)_N)$

β_N: estimated from an empirically-derived algorithm:

$$\beta_N = \overline{\beta_{i,n}}\left[0.8(N'^{-0.98}) + (1-0.8)\right],$$

α_N: $MWDA_N/\beta_N$

$P95_N$: calculated from the gamma distribution parameters α_N and β_N.

The calculation of *r(w/d)* and r*(wet)* are described in Sections 3.4.1.1 and 4.4.1.1 respectively. For both regions, all available stations in the region were used to fit the correlation decay curves used to estimate *r(w/d)* and *r(wet)*.

In order to demonstrate the differences between the possible results when evaluating GCMs against observed station data, the results presented below also show the 'observed' values that might be calculated by:

- the mean parameters calculated for each of the *n* stations (a mean 'point' observation), indicated in **yellow** on figures.
- the parameters of the mean time series of the *n* available stations (an areal observation which may not be representative of the true areal mean), indicated in **green** on figures.

Comparisons are also presented in the form of gamma distributions, fitted to the simulated data and estimated from the observed data values. They show only the distributions of the model data (**red**), the distribution of observed areal values (**blue**). The frequencies indicated by these graphs are scaled by the wet-day probabilities for those rainfall series so that they represent the probability of occurrence of particular precipitation amounts, not the probability conditional upon the occurrence of a wet day. This means that to achieve a good match between the distributions estimated from simulated and observed data requires a good simulation of the number of wet days as well as the distribution of wet-day amounts.

5.4. Model Evaluation over the UK

The 170 station BADC dataset, described in Section 3.2 was again used here for the UK evaluation.

5.4.1. HadCM3

The four 'land' grid boxes covering the UK were analysed from the HadCM3 simulations (Figure 5-1).

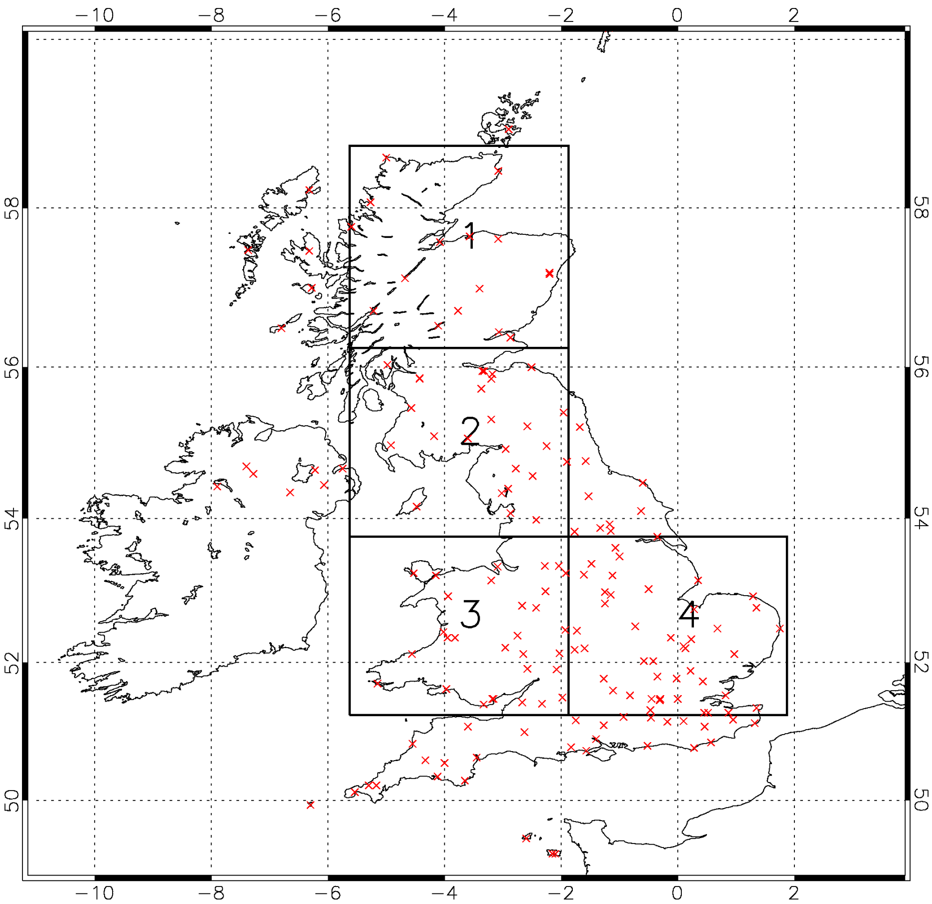

Figure 5-1: HadCM3 UK grid box positions.

5.4.1.1. Mean Daily Rainfall, Dry-Day Frequency and Mean Wet Day Amount

On the whole, HadCM3 fails to reproduce a realistic annual cycle of mean daily rainfall with model runs indicating that JJA is the wettest season, when observations indicate that SON is the wettest (Figure 5-2 toFigure 5-5). SON tends to be too dry relative to the other seasons whilst JJA tends to be too wet. For boxes 1 and 4, the mean daily rainfall is also too high in MAM and DJF, such that annually, these regions are too wet, whilst Boxes 2 and 3 show more accurately simulated DJF and MAM rainfall. This may potentially be part of a geographical trend towards an over-estimation of rainfall in rain shadow regions, which the model is unlikely to be able to capture at this spatial resolution.

The over-estimation of rainfall in summer months (JJA) appears to results from a combination of too many wet days (slightly lower than *P(d)* than observed), and too much rain on those wet days. Mean wet day amount is over estimated by up to 1mm d^{-1}. The under estimation of rainfall in Autumn months (SON) seems to be a result of too many dry days, whilst wet day amounts are simulated more accurately. This is consistent across all boxes.

The differences between model performance for different boxes in Winter and Spring relate to the mean wet day amount. All the boxes demonstrate too many dry days in both these seasons, but for boxes 1 and 4 (Figure 5-2 and Figure 5-5), because the mean wet day amount is around 0.5 mm d-1 too high, the overall mean daily rainfall is also too high. For boxes 2 and 3 (Figure 5-3 and Figure 5-4), the over-estimation of wet-day rainfall amounts is only enough to offset the low wet-day frequency, and the results is that the mean daily rainfall appears to be more accurate.

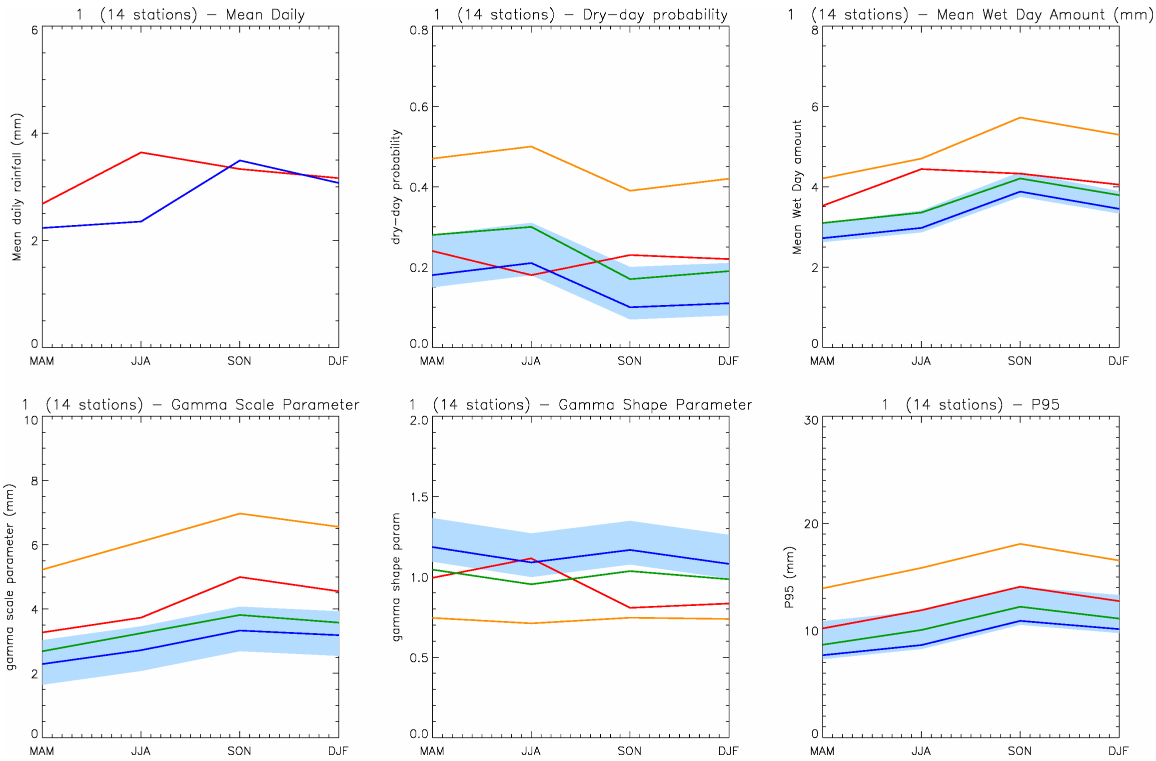

Figure 5-2: Model Evaluation Results for HadCM3 UK Grid box 1. Red=Model, Blue=Observed (best estimate of true areal mean, together with 95% confidence intervals). Also shown are Yellow=average observed 'point' values, Green=observed 'areal' values as arithmetic mean of available station series'.

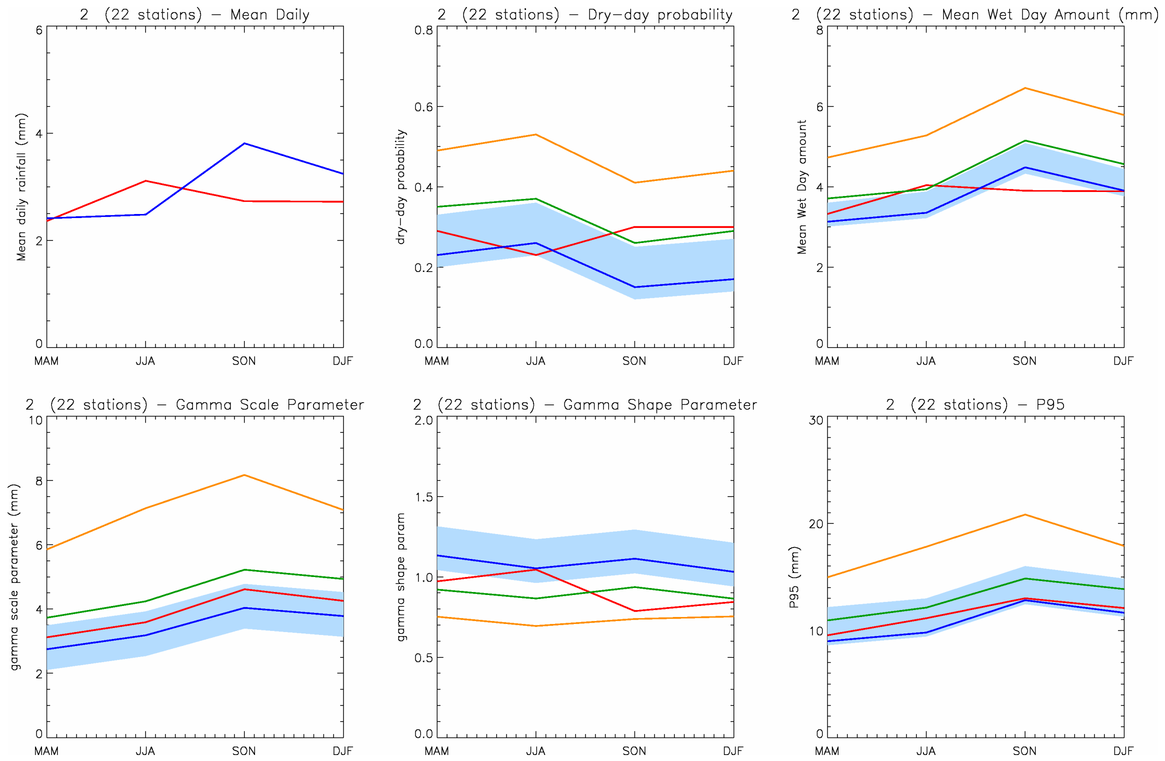

Figure 5-3: Model Evaluation Results for HadCM3 UK Grid box 2.

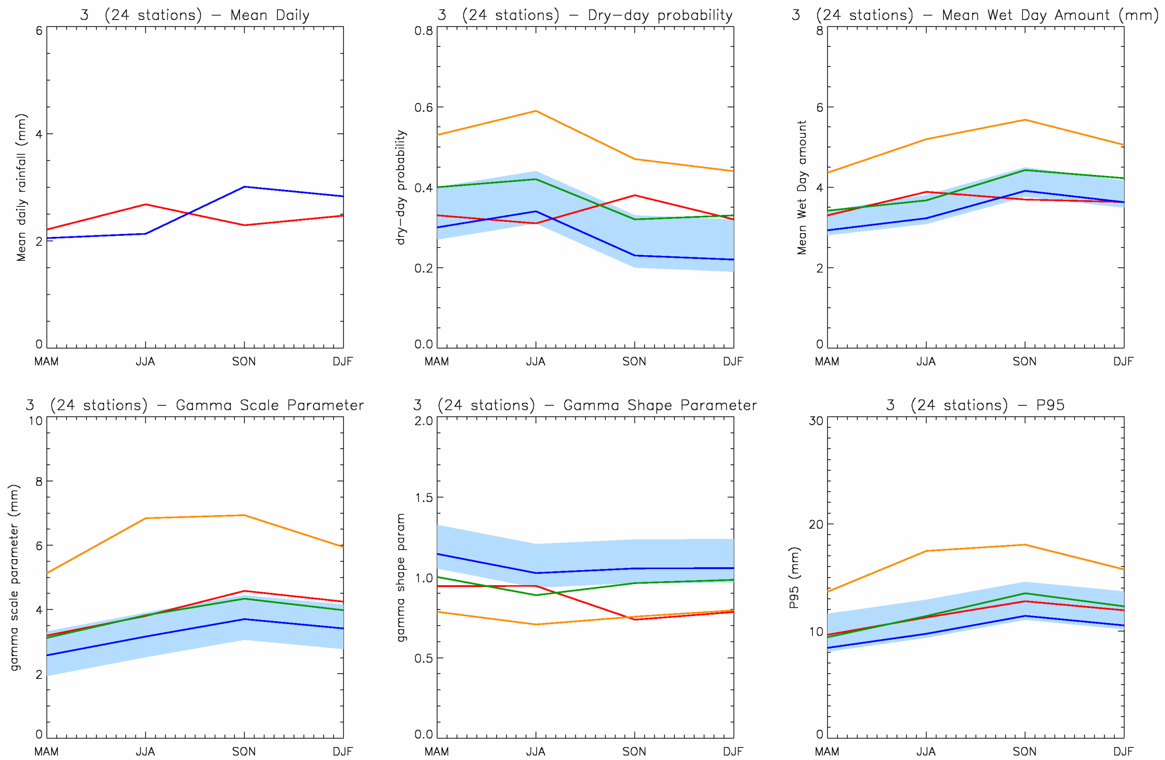

Figure 5-4: Model Evaluation Results for HadCM3 UK Grid box 3.

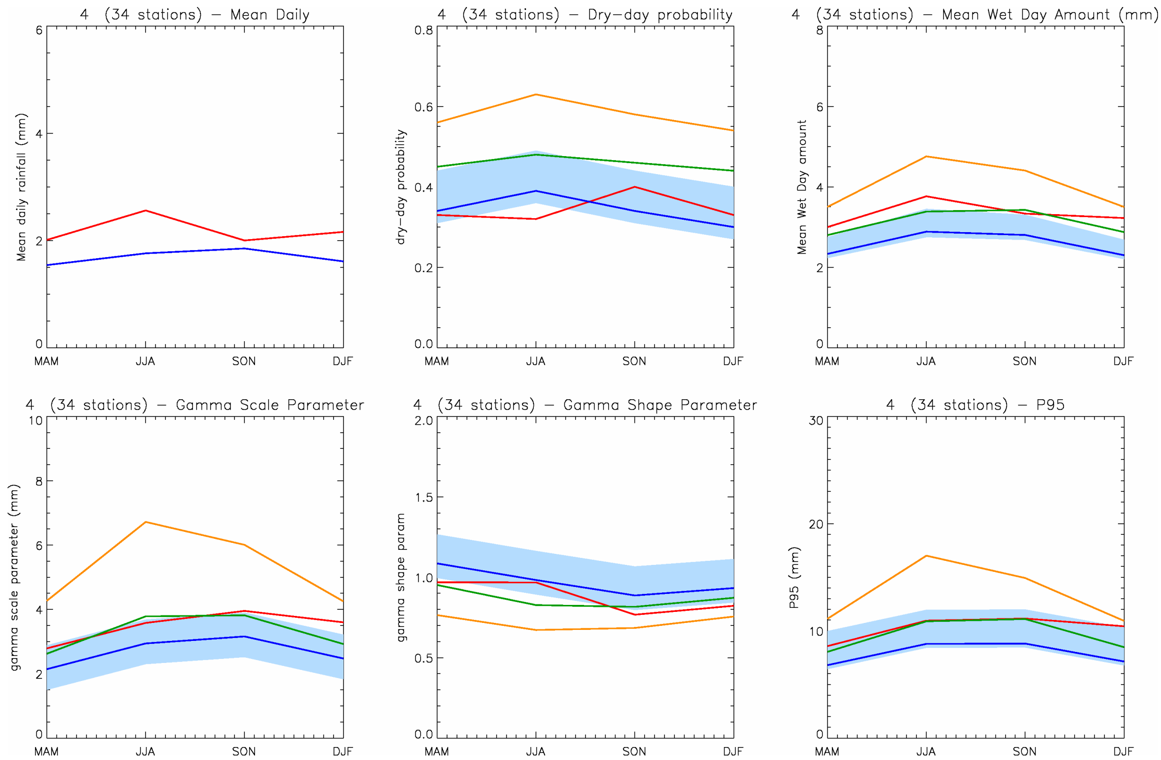

Figure 5-5: Model Evaluation Results for HadCM3 UK Grid box 4.

5.4.1.2. Distribution of Daily Rainfall Amounts

The gamma distribution parameters (Figure 5-2 to Figure 5-5) indicate that the distributions of wet day values simulated by HadCM3 tend to have higher scale parameters than observed, the values of which lie at or beyond the 95% confidence limits in the majority of cases, but a lower shape parameter. This is consistent across seasons SON, DJF and MAM. The seasonal cycle in the scale parameter, representing the range of variability in the daily values, is relatively good, but in the case of the shape parameter, the underestimation is consistent for all seasons except JJA, when this parameter is closer to that observed. The over-estimation of the scale parameter means that the 95th percentile values tend also to be over-estimated, but not enough to exceed the 95% confidence limits.

The effect of these differences in the distribution is shown by comparing the distribution probability density functions, (Figure 5-6). In summer (JJA), the shape parameter is relatively well simulated and the distributions are of similar shape for lower rainfall values. For the remaining three seasons, particularly SON, the shape parameter is a little too low, demonstrating that the proportion of ‘light’ or ‘drizzly’ rainy days is higher in the model than in observations. This is the most marked in SON.

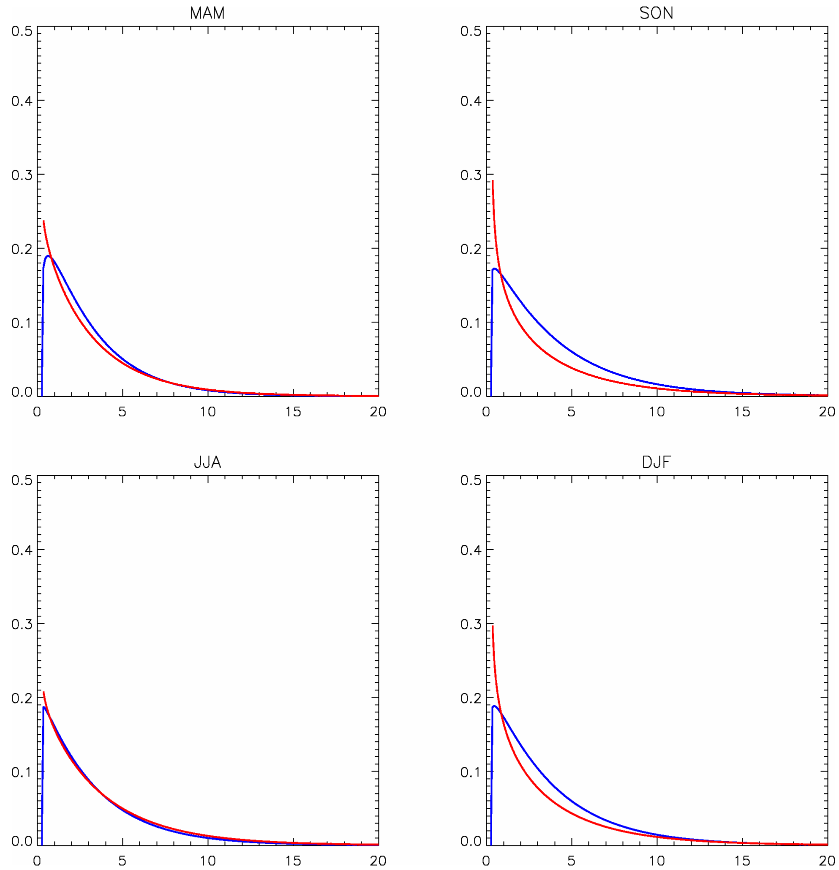

Figure 5-6: Gamma probability density functions fitted to HadCM3 Box 3 wet-day amounts, multiplied by box 3 wet-day probability.

5.4.1.3. Summary

The rainfall amounts, and parameter values reflecting the characteristics of that rainfall, are realistic in terms of their magnitude in HadCM3 simulations, but the seasonal cycle is not well simulated. Overall, the model tends to be a too wet in summer and too dry in autumn.

The distribution of daily values simulated is generally good compared to the observed, but summer values tend too be slightly too frequently 'heavy', causing the season to be a little too wet, whilst Autumn (SON) rainfall tends to be too drizzly, with too few heavier rainfall days.

The green and orange curves shown in Figure 5-2 to Figure 5-5 give an indication of the different conclusions that may be drawn about model performance if these different 'observed' records were used. All four grid boxes examined here have relatively good station coverage (> 14 stations), and so whilst the difference between the yellow (the average point value) and blue (estimated areal average value) is very distinct, the difference between the green (average of available stations) and the blue is less substantial. The green line often lies within, or very close to, the uncertainty margins of the 'best' estimate' of the areal mean (light blue). The difference between these two sets of values does, however, still affect the results of the evaluation. The values of the evaluated parameters for the model simulated rainfall are in many cases closer to the green line than the blue, particularly for the gamma scale parameter and 95th percentile. This suggests that the model performs slightly less well than might be found if a simple average of available stations were used for a quantitative evaluation. The direction of error is also depends on which estimate of areal rainfall is used in some cases (e.g. The gamma scale parameter for box 2 - Figure 5-3).

5.4.2. CCC CGCM3

Five grid boxes were analysed for the UK from the CGCM3 simulations. An ensemble of 4 runs was available and all are used in this evaluation, giving an indication of the variation in model results caused by differing initial conditions.

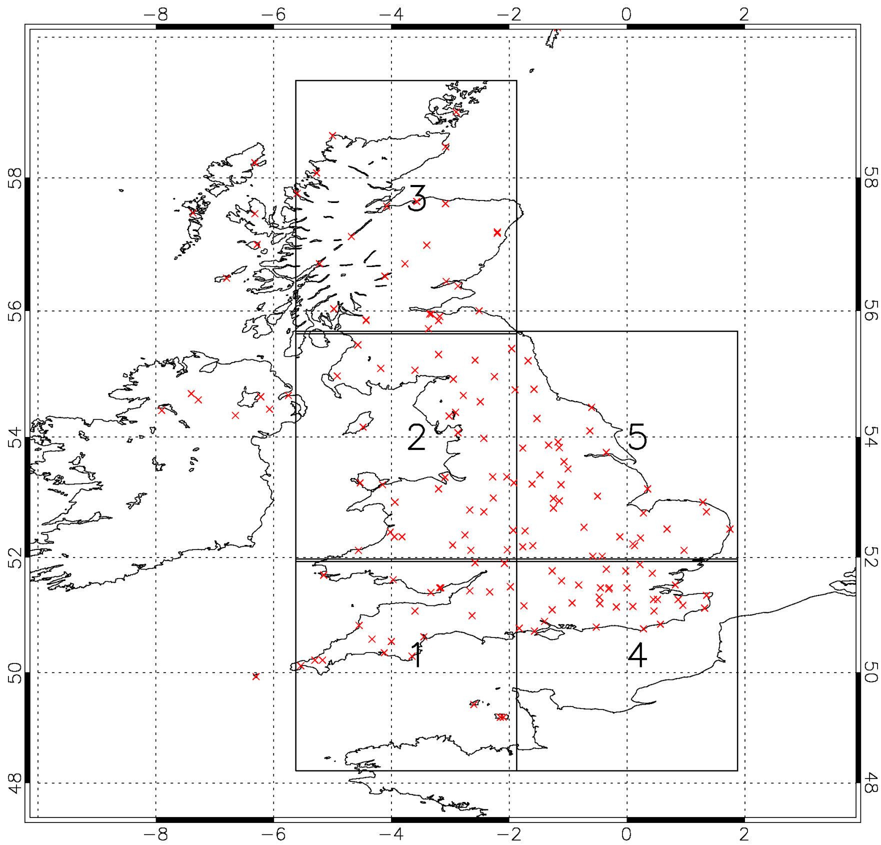

Figure 5-7: CGCM3 UK grid box positions

5.4.2.1. *Mean Daily rainfall, Dry-Day Frequency and Mean Wet-Day Amount*

The mean daily rainfall in the three western UK grid boxes, boxes 1, 2 and 3, is underestimated by 0.5 to 1mm, in a relatively consistent pattern across the seasons, so that the annual cycle is well simulated. The two eastern boxes, however, have more realistic overall rainfall amounts, which are even slightly over-estimated for box 5.

The rainfall simulated by CGCM3 for the UK is characterized by consistent underestimation of the dry-day probability (by as much as 0.2) and mean wet day amount (by as much as 1.5mm per day), the margin of which are large enough to lie outside the error margins of the observed estimates in almost all cases. The annual cycles in these values are also relatively well preserved, although JJA dry-day probability tends to be underestimated by a larger magnitude than the other seasons.

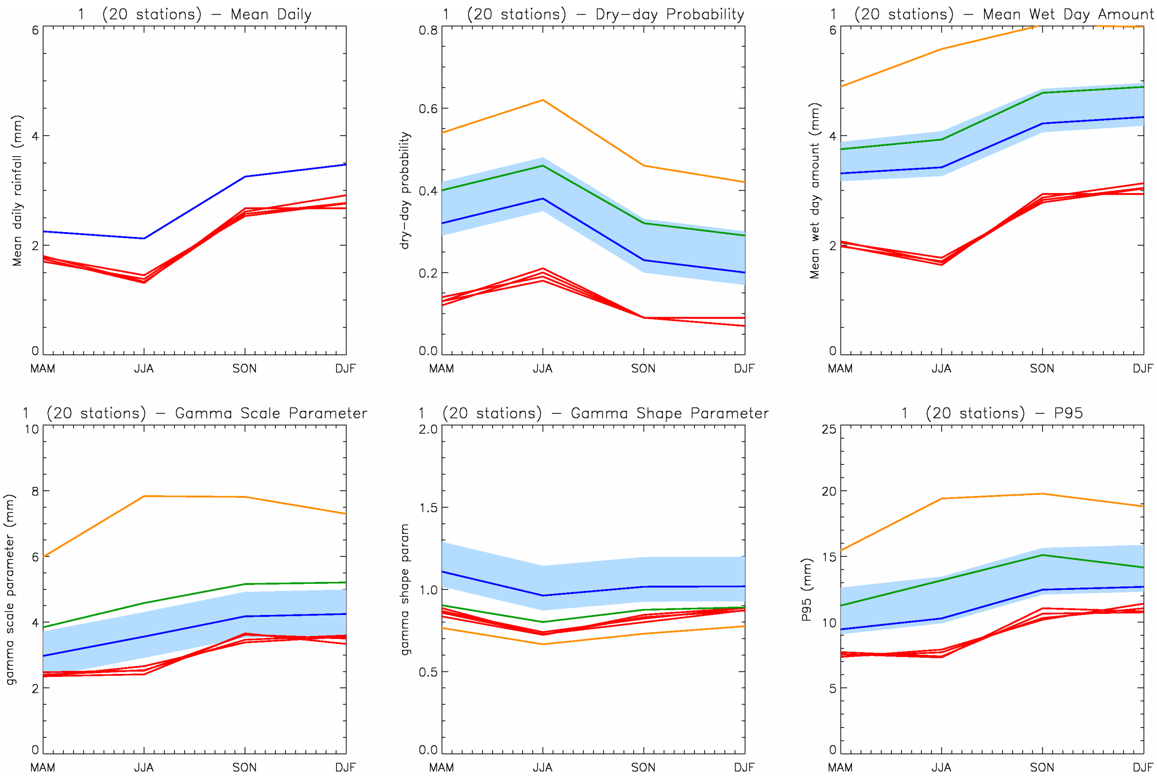

Figure 5-8: Model Evaluation Results for CGCM3 UK grid box 1. Red=Model, Blue=Observed (best estimate of true areal mean, together with 95% confidence intervals). Also shown are Yellow=average observed 'point' values, Green=observed 'areal' values as arithmetic mean of available station series'.

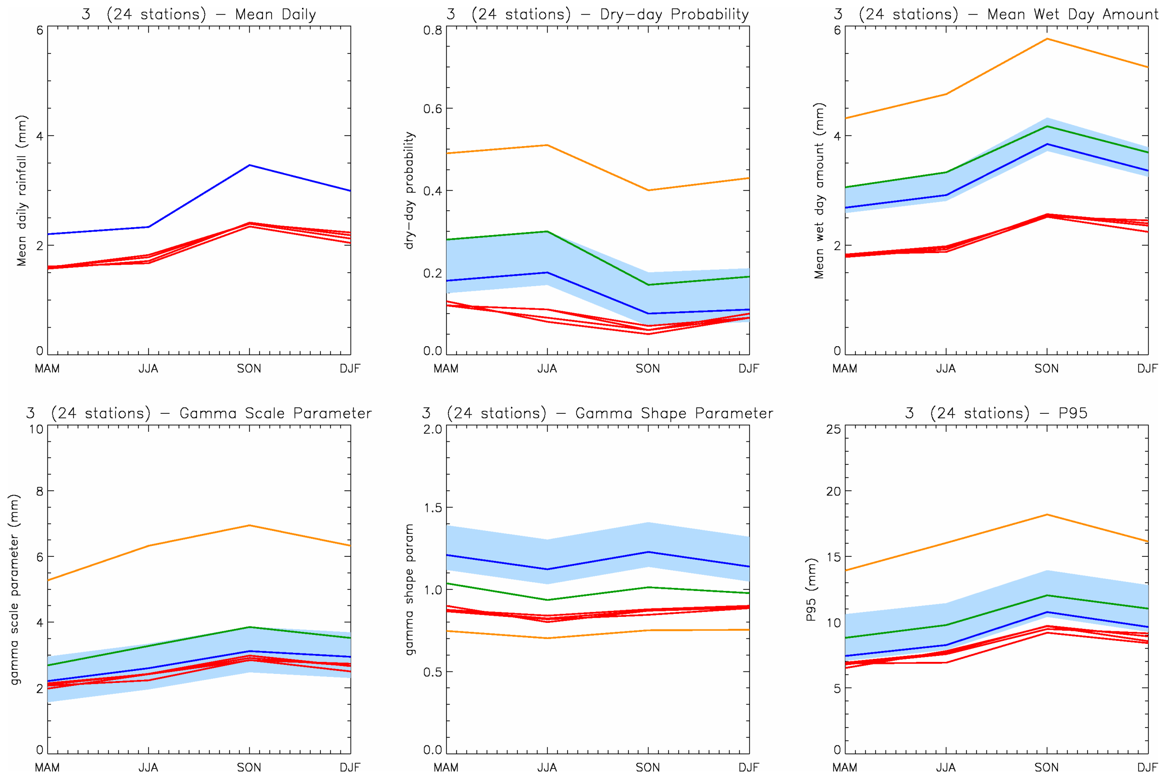

Figure 5-9: Model evaluation results for CGCM3 UK grid box 2.

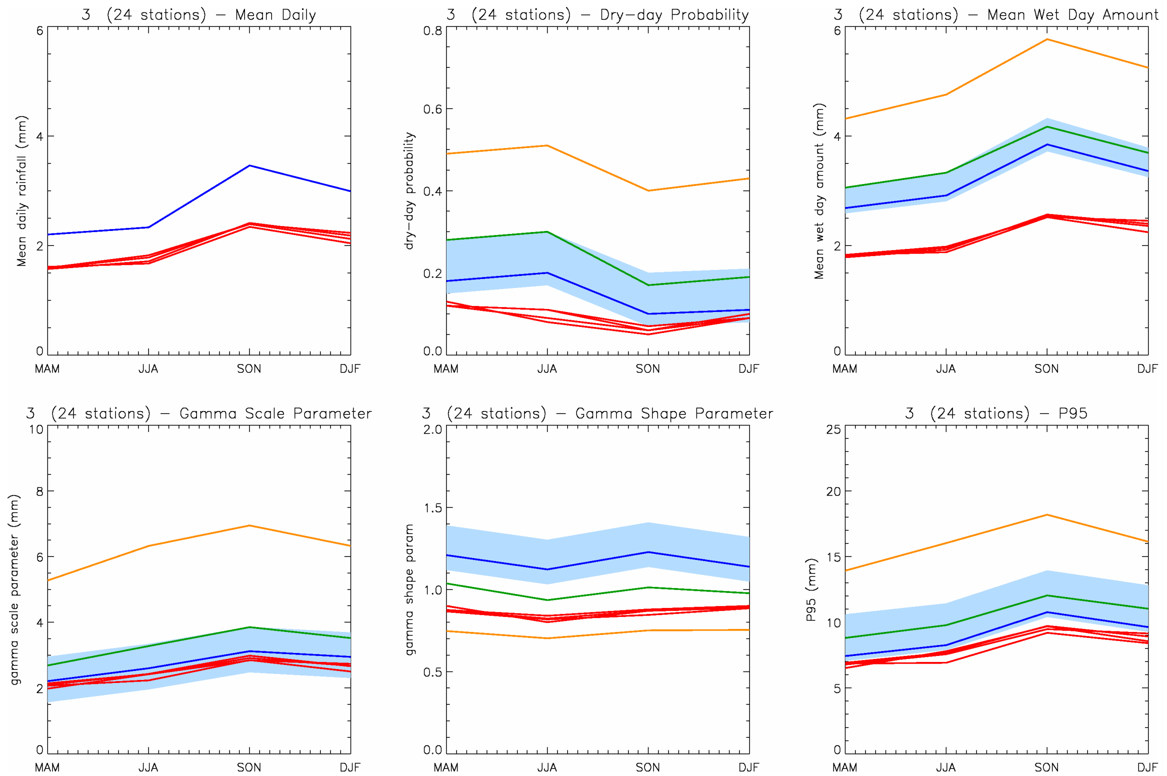

Figure 5-10: Model evaluation results for CGCM3 UK grid box 3.

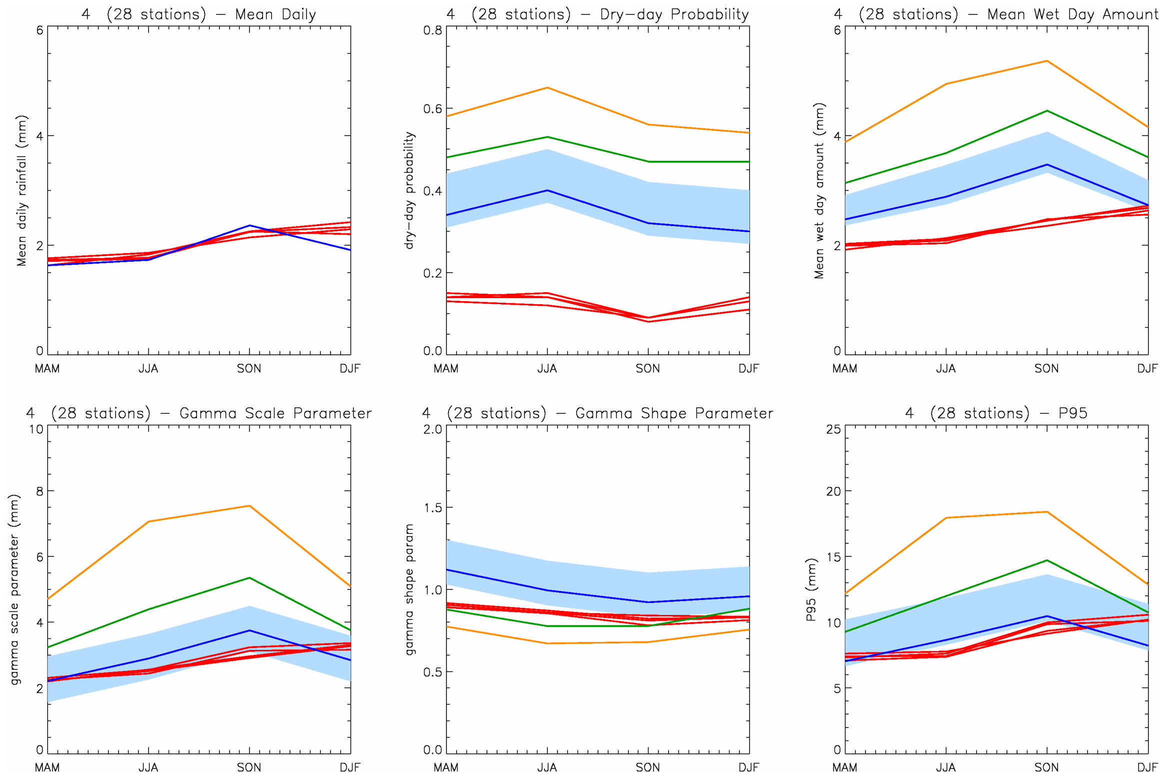

Figure 5-11: Model evaluation results for CGCM3 UK grid box 4.

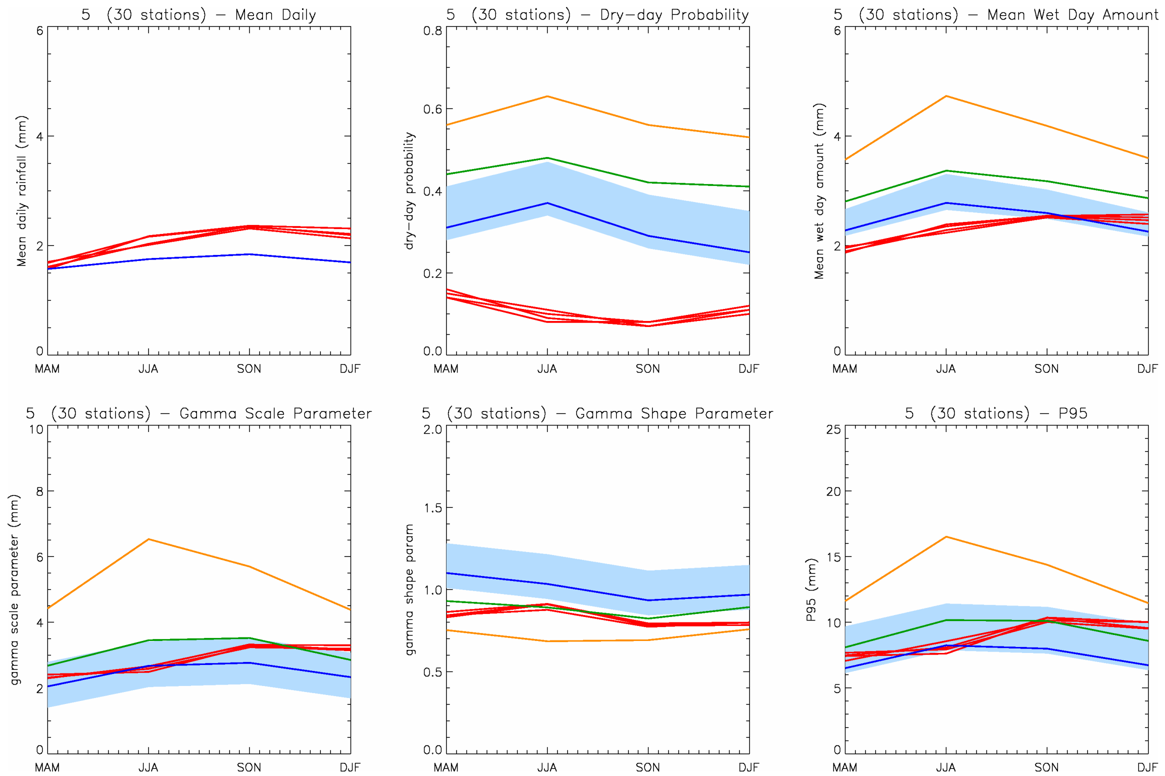

Figure 5-12: Model evaluation results for CGCM3 UK grid box 5.

5.4.2.2. Distribution of Wet-Day Values

The simulated rainfall is distributed with a reasonably accurate scale parameter for most regions of the UK. This value rarely falls beyond the confidence limits of the observed scale parameter.

The shape parameter of the simulated distribution, however, is consistently lower than the observed, indicating a consistently higher proportion of 'drizzly' or 'light' rainfall day. An example of the modeled and observed distributions for grid box 3 is shown in Figure 5-13. According to the model, a 0.5-1.5mm event occurs on around 40-50% of days, whilst the observations suggest this should be more like 20%. This is most marked in the western grid boxes, but does not vary consistently according to season. The model also appears to underestimate slightly the frequency of moderate rainfall days (2mm or more) in the western grid boxes, which causes the total rainfall to be too low.

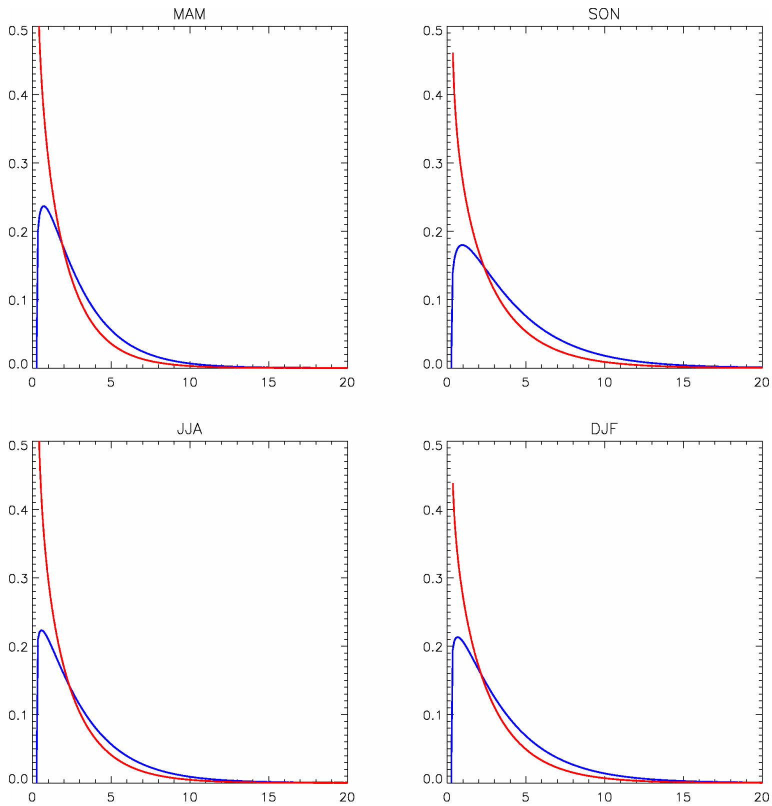

Figure 5-13: Gamma probability density functions fitted to CGCM3 grid-box 3 wet-day amounts, multiplied by Box 3 wet-day probability.

5.4.2.3. Summary

The rainfall simulated by CGCM3 for the UK tends to be too 'drizzly' in nature, with rain occurring on too many days, but generally in smaller amounts than are observed. Rainfall events in the range 0-5-1.5mm occur on 40-50% of days in the model rather than around 20% of days seen in the observation-based estimates.

For all of the statistical parameters evaluated except the gamma shape parameter, the model values are closer to the 'best guess' estimate of areal rainfall (blue line) than to the simple station average (green). The direction of model error remains the same regardless of which approach is used in almost all cases, but the magnitude of the error is smaller when the estimate of 'true' areal mean is used as the observed benchmark.

5.4.3. PCM

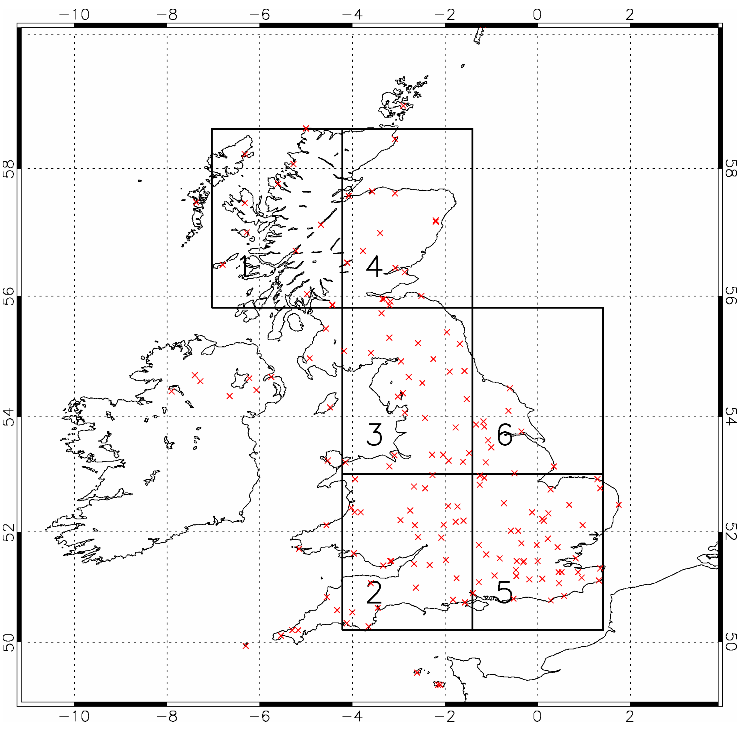

Figure 5-14: PCM UK grid box positions

5.4.3.1. Mean Daily Rainfall, Dry-Day Frequency and Mean Wet-Day Amount

Like CGCM3 simulations, the precipitation for UK boxes from PCM consistently include too few dry-days, with error magnitudes of up to 0.3, which far exceeds the confidence limits of the observed series (Figure 5-15 to Figure 5-20).

The realism of the simulated mean wet day amounts varies from region to region, but for most grid boxes is more realistic than those of CGCM; only box 1 shows a substantial under simulation of wet day rainfall. Most boxes have values of an appropriate magnitude for all seasons except summer, for which rainfalls tend to be smaller than those which are observed.

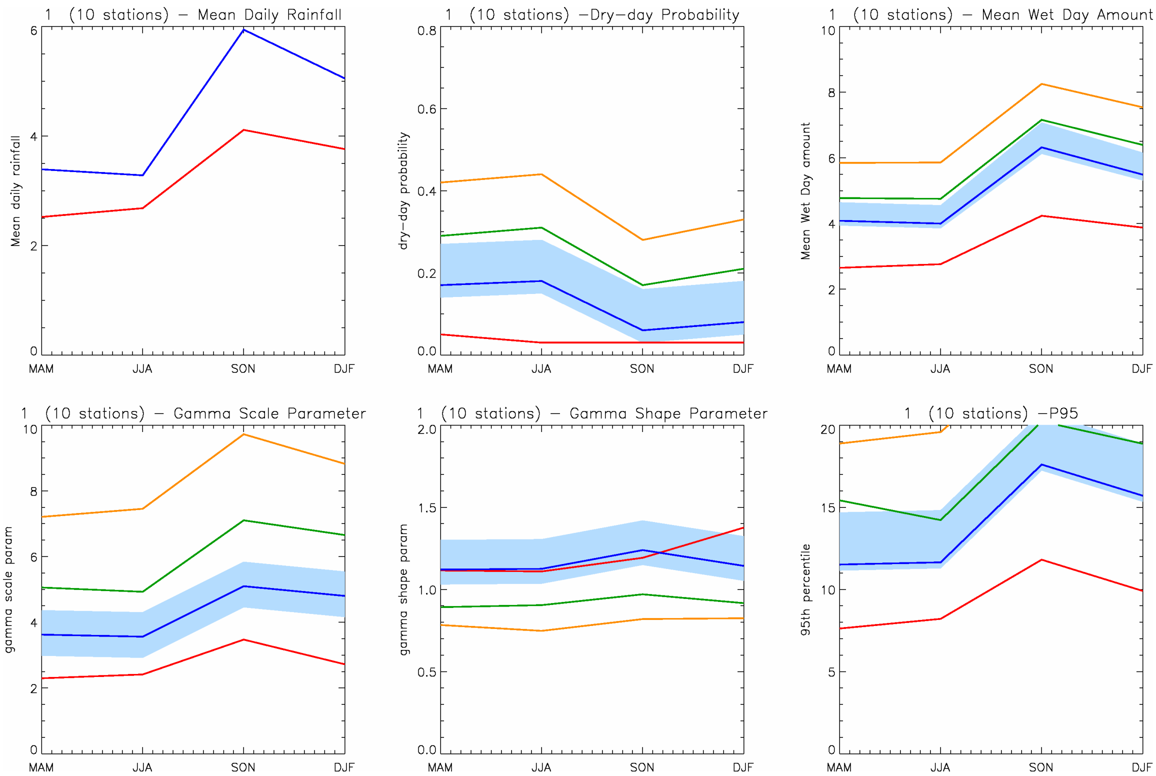

Figure 5-15: Model evaluation results for PCM UK grid box 1. Red=Model, Blue=Observed (best estimate of true areal mean, together with 95% confidence intervals). Also shown are Yellow=average observed 'point' values, Green=observed 'areal' values as arithmetic mean of available station series'.

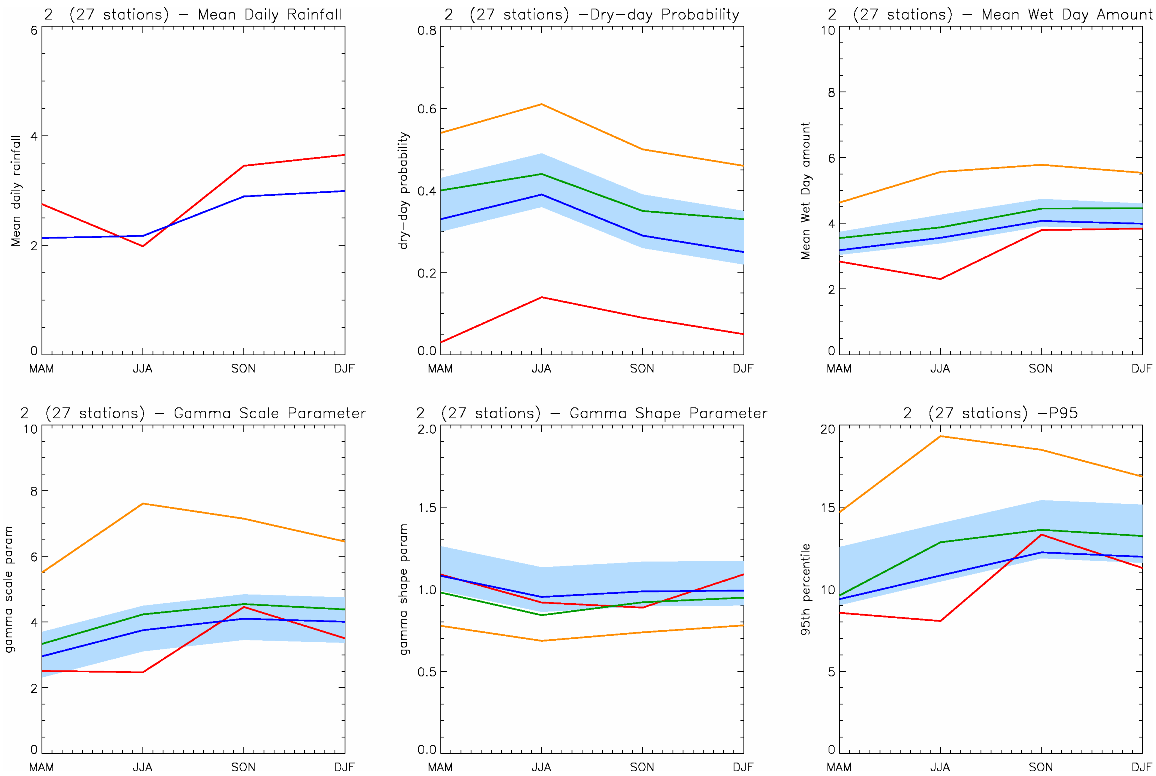

Figure 5-16: Model evaluation results for PCM UK grid box 2.

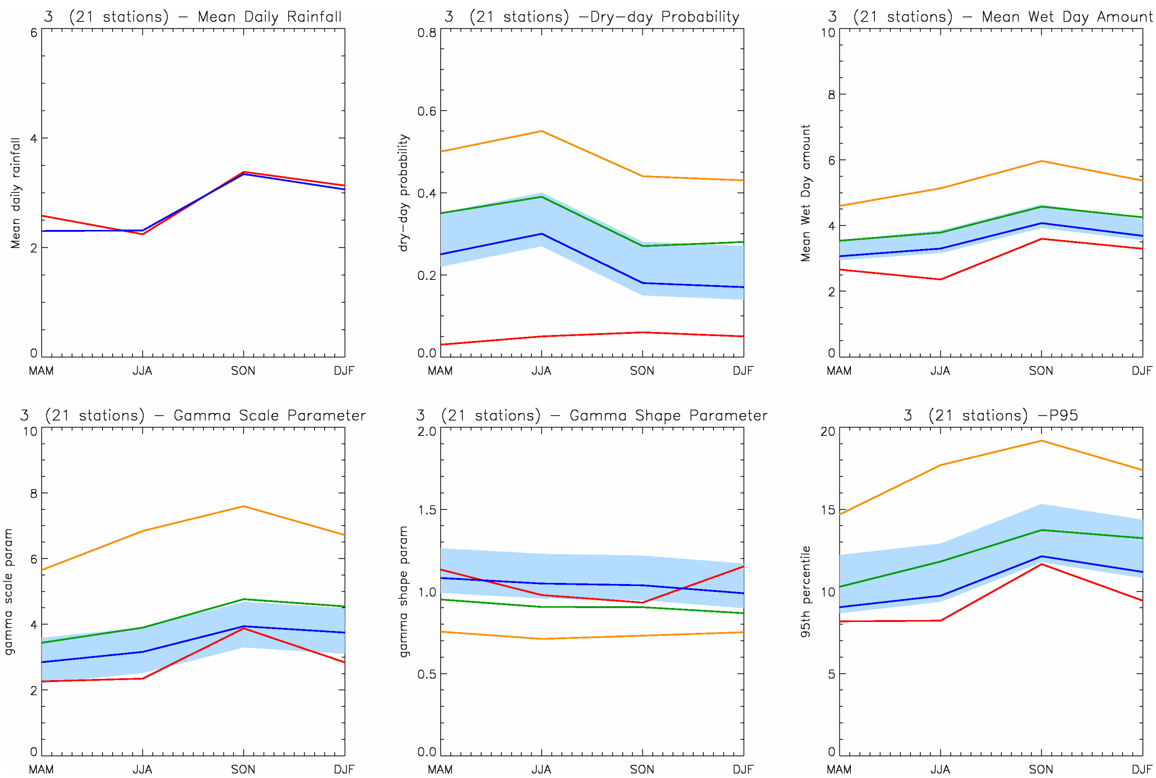

Figure 5-17: Model evaluation results for PCM UK grid box 3.

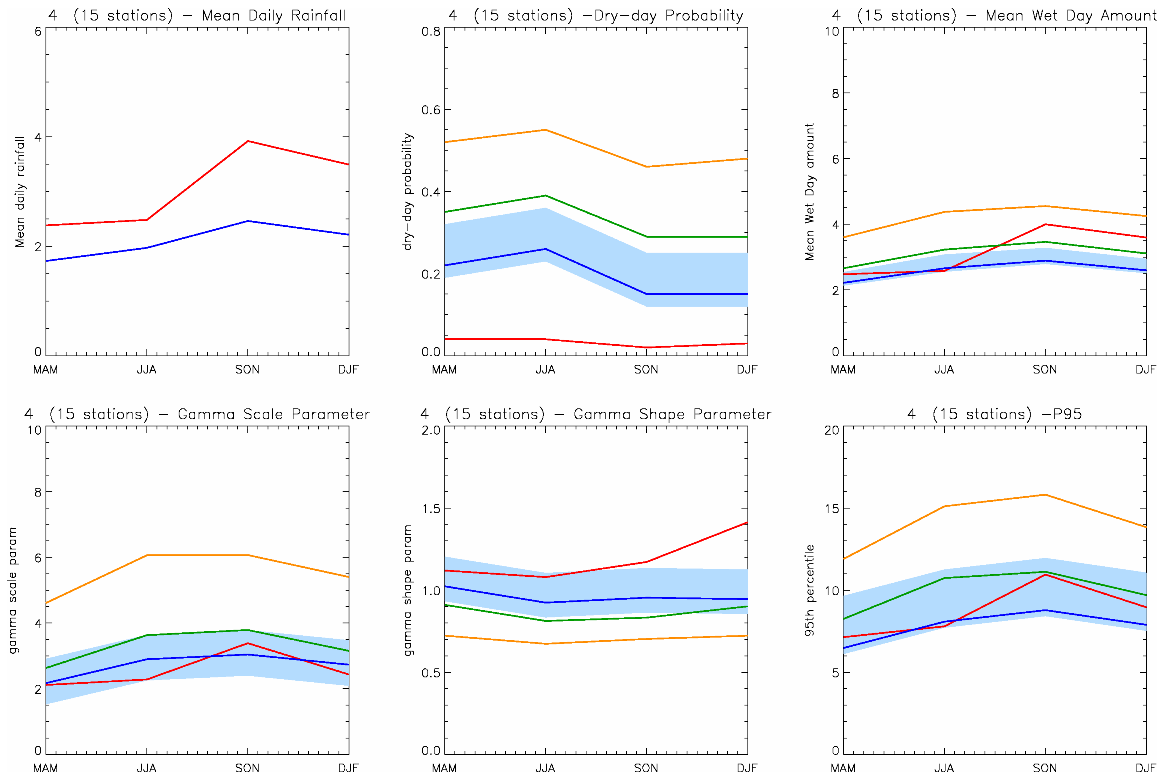

Figure 5-18: Model evaluation results for PCM UK grid box 4.

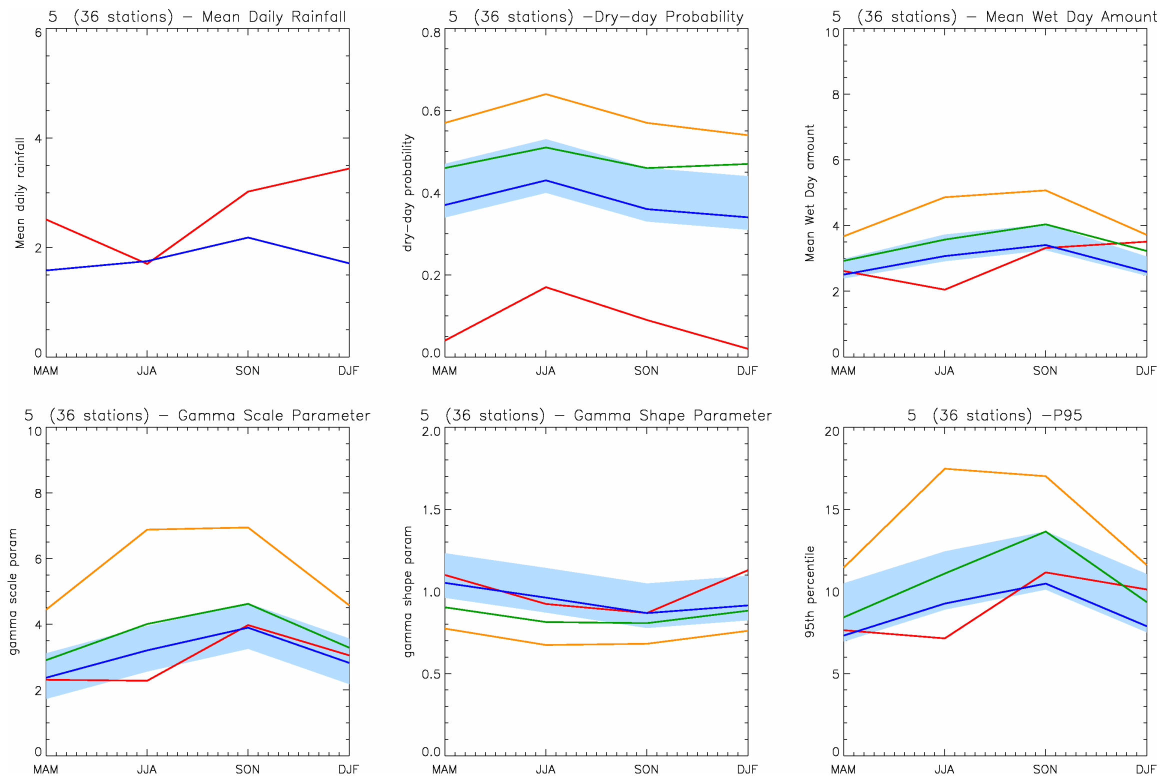

Figure 5-19: Model evaluation results for PCM UK grid box 5.

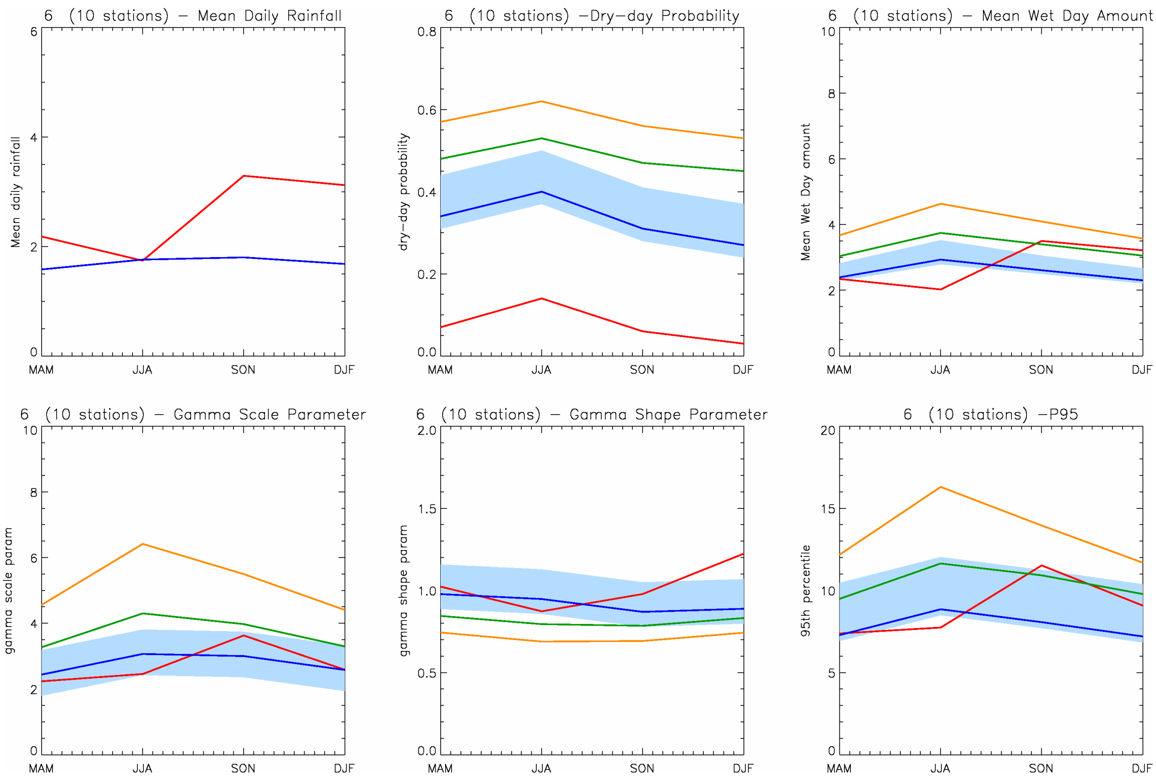

Figure 5-20: Model evaluation results for PCM UK grid box 6.

5.4.3.2. Distribution of Wet-Day Values

The distributions of wet-day values by this model appear to be relatively reliable (Figure 5-15 to Figure 5-20). The shape parameter is generally of an appropriate magnitude, with modelled values lying within the confidence limits of the observed for most boxes and seasons except winter when the shape parameter is consistently over estimated by around 0.25. The scale parameter also lies within the confidence limits of the observed for most seasons and grid boxes There is a tendency for this value to be underestimated in summer and over-estimated in autumn, resulting in the respective under and over-estimation of extreme values, illustrated by the 95th percentile of the wet-day amounts.

The distributions of values are therefore reasonably well simulated by PCM. The probability density functions (Figure 5-21) show that rainfall events of all magnitudes are too frequent. Only the season JJA experiences proportionally too much light rainfall.

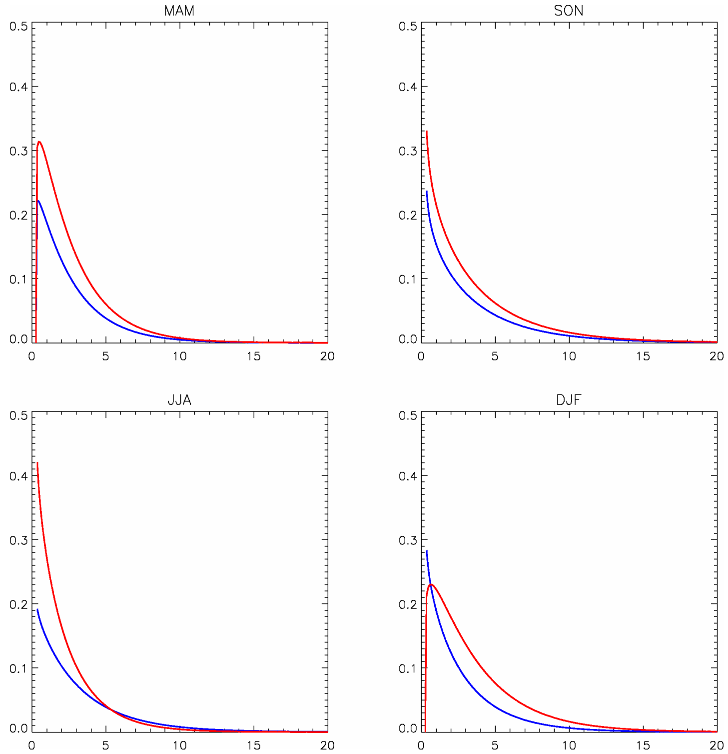

Figure 5-21: Gamma probability density functions fitted to PCM grid-box 5 wet-day amounts, multiplied by Box 5 wet-day probability.

5.4.3.3. Summary

Rainfall simulated by PCM for the UK also appears to occur on too many days. However, unlike CGCM3, which tends to simulate a disproportionate number of 'drizzly' days, the distribution of wet-day values from PCM is more similar to the observed, and the tendency seems to be that wet days of all magnitudes are a little too frequent. This is the case for all seasons but summer, when the total rainfall amount appears to be more accurate because the over-estimation of wet-day frequency is offset by under-estimated wet-day amounts, and the distribution is biased towards too many light/drizzly days.

Similarly to the CGCM3 evaluation, the model performance is generally better when compared with the estimate of 'true' areal mean rainfall parameters (blue line) than when a simple average of available stations is used (green lines). The use of the 'best guess of the true areal mean' instead of the simple average of available stations does not affect the direction of the model error, but the magnitude of the error is less when the former is taken to be the 'observed'. Use of the simple average of available stations may therefore result in an under-estimation of the model's ability to reproduce the characteristics of daily rainfall in this region.

5.5. Model Evaluation over South Africa

5.5.1. South African Data and Methods

A set of daily rainfall observation stations from northern South Africa, acquired from the South African Weather Bureau, are used. These data generally have nearly complete temporal coverage from 1961-2000. The station distributions in relation to the grids of the three GCMs are shown in Figure 5-25, Figure 5-28 and Figure 5-31.

The methods used to estimate the 'true' areal observed precipitation parameter values are the same as those used for the UK (Section 5.3), which were developed in chapters 3 and 4. These methods require values of the mean inter-station 'correlation', for correlation values *r(wet)* (the correlation between wet-days only) and *r(w/d)* (a measure of correlation between wet and dry day occurrence, but not amount). This is calculated for each grid box using correlation decay curves fitted for every station within each grid box, using all the other available stations (not only those contained within that particular grid box). The parameters for the curves are then averaged across the stations within a grid box to give a grid-box average decay curve which can be used to estimate the mean inter-station correlation using the distribution of separation distances that a very large (or infinite) number of station pairs would give (see section 3.4.1.1 for further details of this method) .

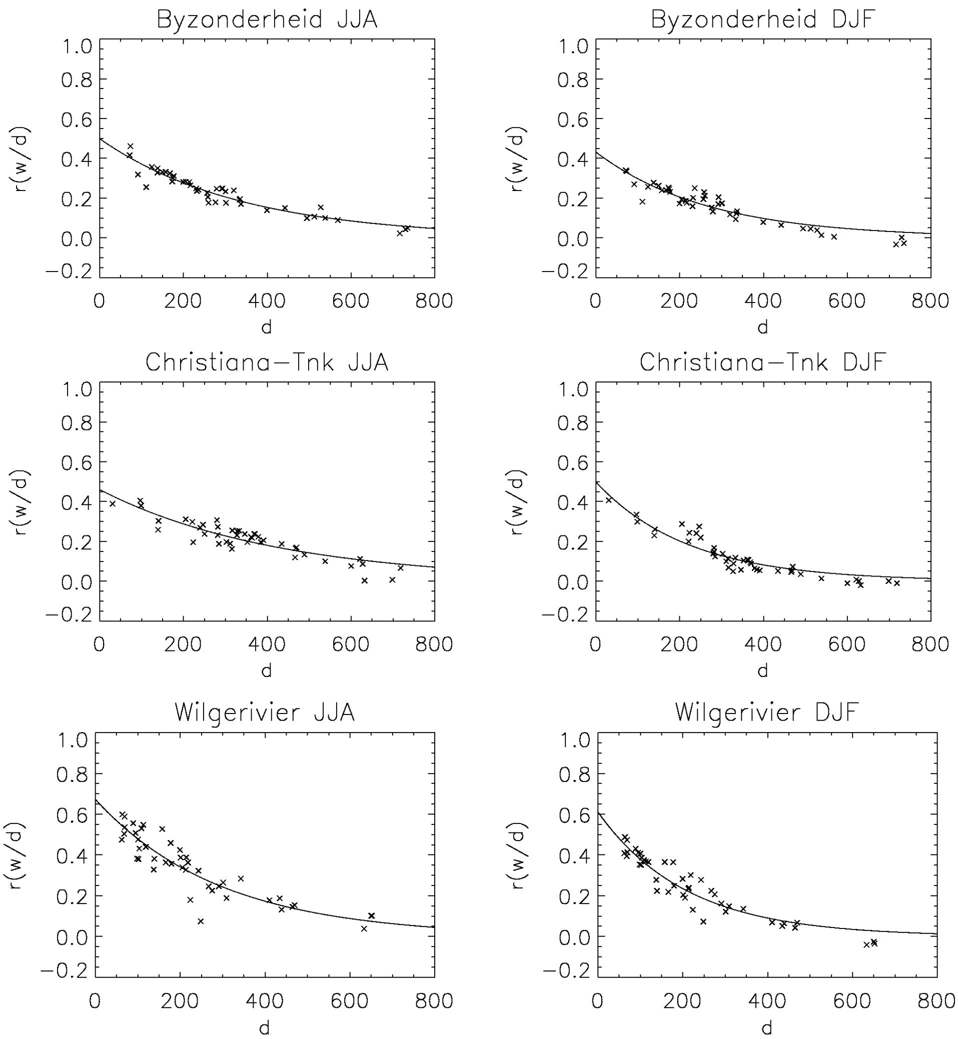

Figure 5-22: Examples of correlation decay curves for *r(w/d),* with curves fitted according to Equation 5-1.

The only difference between the methodologies for the South African data compared to that of the UK data is the function used to fit decay curves. For the UK stations, the mean inter-station correlation, based on *r(wet)* and *r(w/d)* were calculated using correlation decay curves, where the decrease in correlation with increasing separation distance is fitted with the exponential function:

$$r = ae^{-bd}$$

Equation 5-1

In order to fit correlation decay curves for the South African stations, a third parameter, *c*, is required because the values of *r(wet)* fall below zero. The function becomes:

$$r = c + ae^{-bd}$$

Equation 5-2

This adjustment was required for the South African stations as the correlation declines to below zero more rapidly as separation distance *(d)* increases than for stations in the UK. The poor fit of the original function (Equation 5-1) at the large separation distances for the UK does not cause a problem because only separation distances up to around 400km are used for the model evaluation.

The simpler function (Equation 5-1) is still used for *r(w/d)* correlation calculations, which do not decay to values below zero. Some examples of these fitted curves are shown in Figure 5-22 and Figure 5-24.

A problem also arises in the calculation of the 95% confidence limits attached to the estimates of dry-day probability and mean wet day amount. The confidence limits for $P(d)_N$ are derived empirically in Section 3.3.5 and are +0.10 and -0.03. However, where dry-day probability is very high in JJA in South Africa, this approach fails because the upper limit for $P(d)_N$ becomes greater than 1 in several cases. This can be resolved by resetting the upper value to 1 where a larger value occurs.

The mean wet-day amount for the estimated ‘true’ areal mean is calculated via the mean daily rainfall and the estimated areal dry-day probability (Equation 5-3).

$$MWDA_N = \frac{MD_N}{1 - P(d)_N}$$

Equation 5-3

The empirically derived 95% confidence limit values of dry-day probability ($P(d)_N$) of +0.10 and -0.03 are therefore used to calculate similar limits for the mean wet-day amount. However, when the upper limit of $P(d)_N = 1$, Equation 5-3 fails because it causes the denominator of this fraction to become zero. This could be partially resolved by re-setting the $P(d)_N$ value to a value to 0.99 rather than 1, which gives a real, but very high upper limit of the MWDA as mean daily rainfall is divided by 0.01. Instead, the upper limit in such cases is replaced with the mean point value of the dry-day probability or MWDA (yellow line – see Figure 5-23). The dry-day probability in a series which is a mean of any number of individual series' cannot be greater than the mean dry-day probability of all the individual series' – it will always be lower. In this application, this remains valid as long as the assumption that the *n* available stations are representative of every point in the grid box (this is discussed in 3.4.1.2).

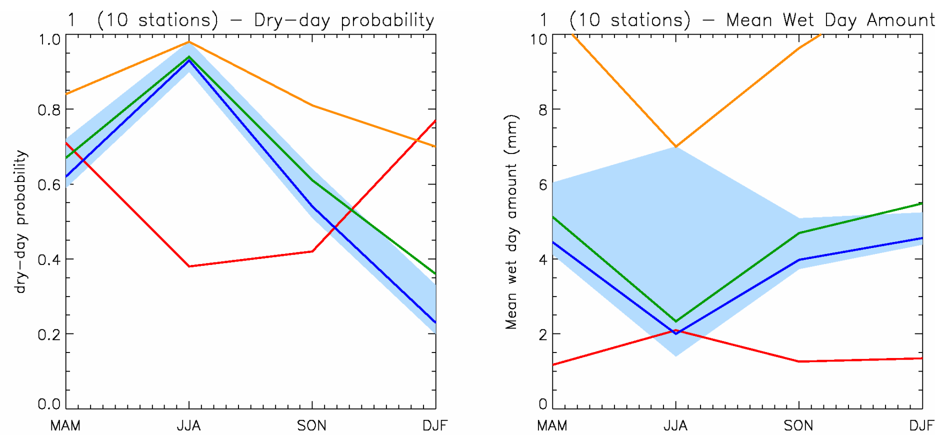

Figure 5-23: Example of confidence limits for South African dry-day probability and mean wet-day amount for areal rainfall estimates. Red=Model, Blue=Observed (best estimate of true areal mean, together with 95% confidence intervals). Also shown are Yellow=average observed 'point' values, Green=observed 'areal' values as arithmetic mean of available station series'.

We must also bear in mind that exploration in Chapter 4 (Section 4.3.4) of the uncertainty in the estimates of the gamma parameters for the *n* station mean, the methods were found to be unreliable in some regions and seasons that are particularly dry (i.e. when mean daily rainfall is less than 0.3mm). This situation arises in the dry season (JJA) of South Africa, and the estimations of 'true' areal mean gamma parameters in these cases must be interpreted with caution.

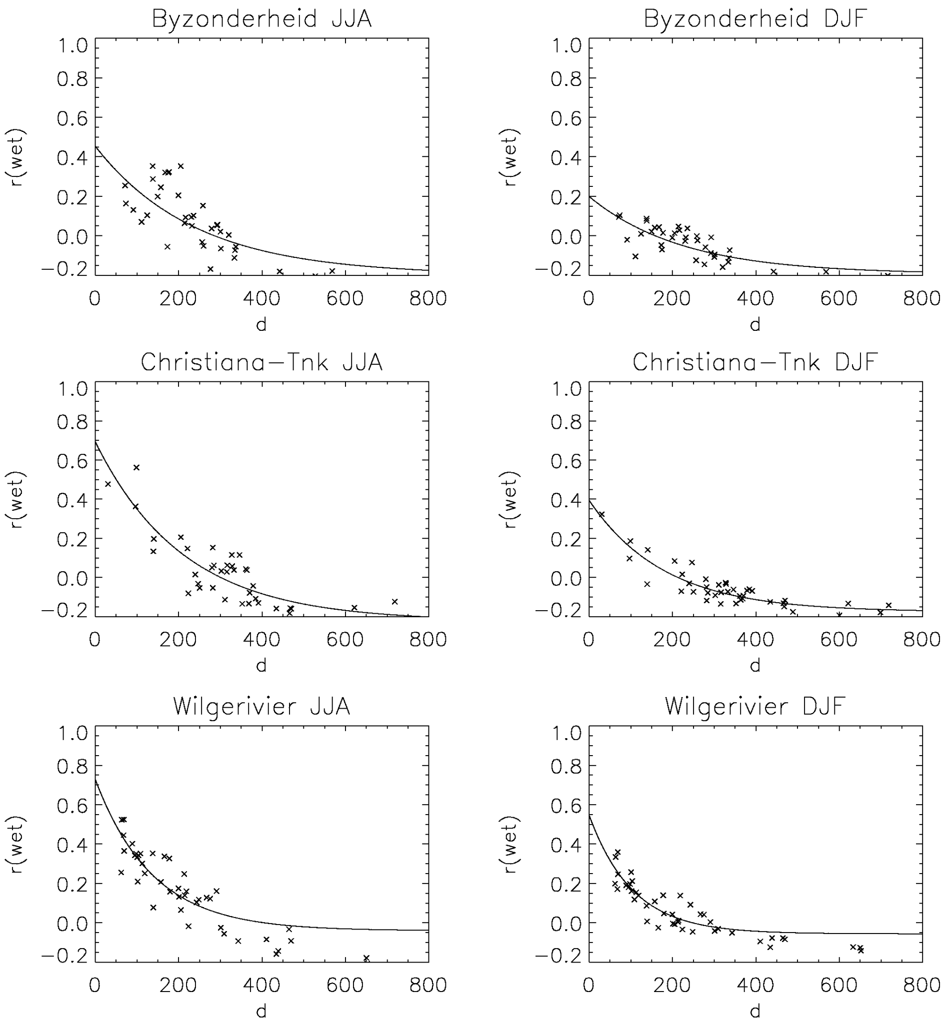

Figure 5-24: Examples of correlation decay curves for *r(wet),* with curves fitted according to Equation 5-2.

5.5.2. HadCM3

Four grid boxes from the HadCM3 dataset were evaluated, the positions of which are shown in Figure 5-25. The results for one of these boxes (box 1) is shown here as an example (Figure 5-26 and Figure 5-27).

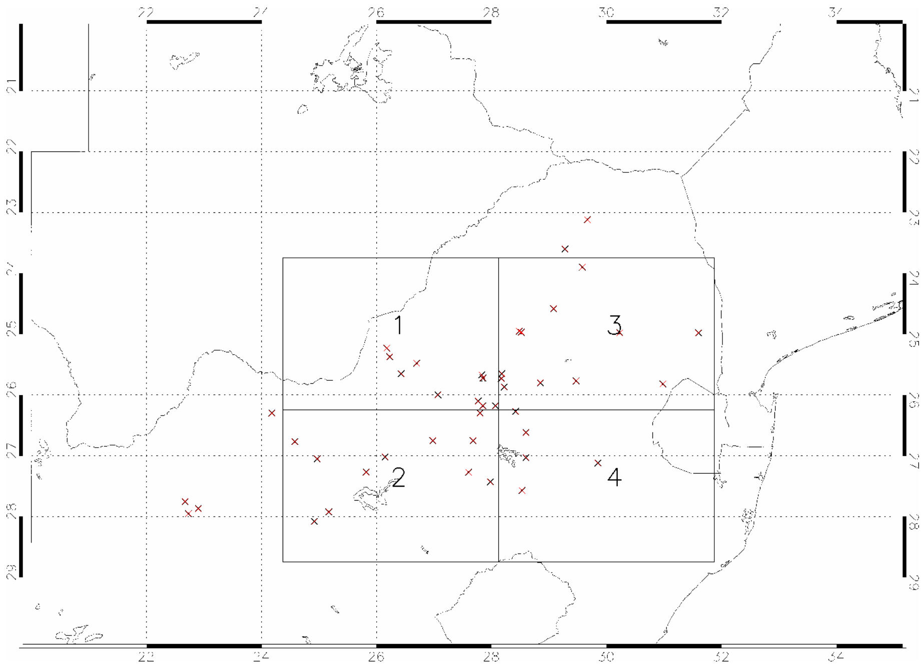

Figure 5-25: HadCM3 South African grid box positions and rain gauge locations.

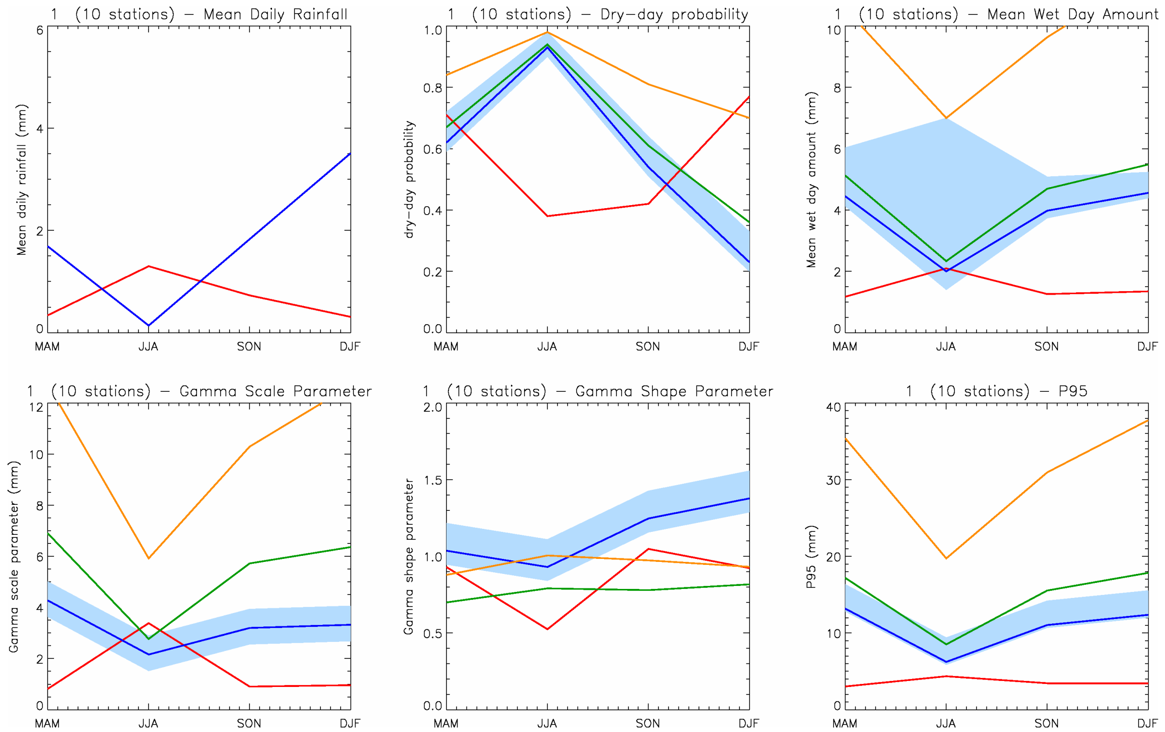

Figure 5-26: Model Evaluation Results for HadCM3 SA grid box 1. Red=Model, Blue=Observed (best estimate of true areal mean, together with 95% confidence intervals). Also shown are Yellow=average observed 'point' values, Green=observed 'areal' values as arithmetic mean of available station series'.

It was noted in the evaluations of UK grid boxes that HadCM3 struggles to reproduce the seasonal patterns of rainfall amount, and this appears to be the case also for South African rainfall (for this grid box and also for the other 3 boxes which are not shown). The overall mean rainfall is consistently too low in all seasons except but JJA, and does not reflect the dry and wet seasons (JJA and DJF respectively) which are distinctive in the observed series.

This poor seasonal representation affects the other characteristics of the rainfall as well as the overall amount. Dry-day frequency is substantially underestimated in JJA and SON such that the JJA dry season is too wet, but overestimated in the other seasons such that they are generally too dry. Mean wet day amount is substantially underestimated (by around 2mm in MAM, DJF and SON, and 0.5mm in JJA).

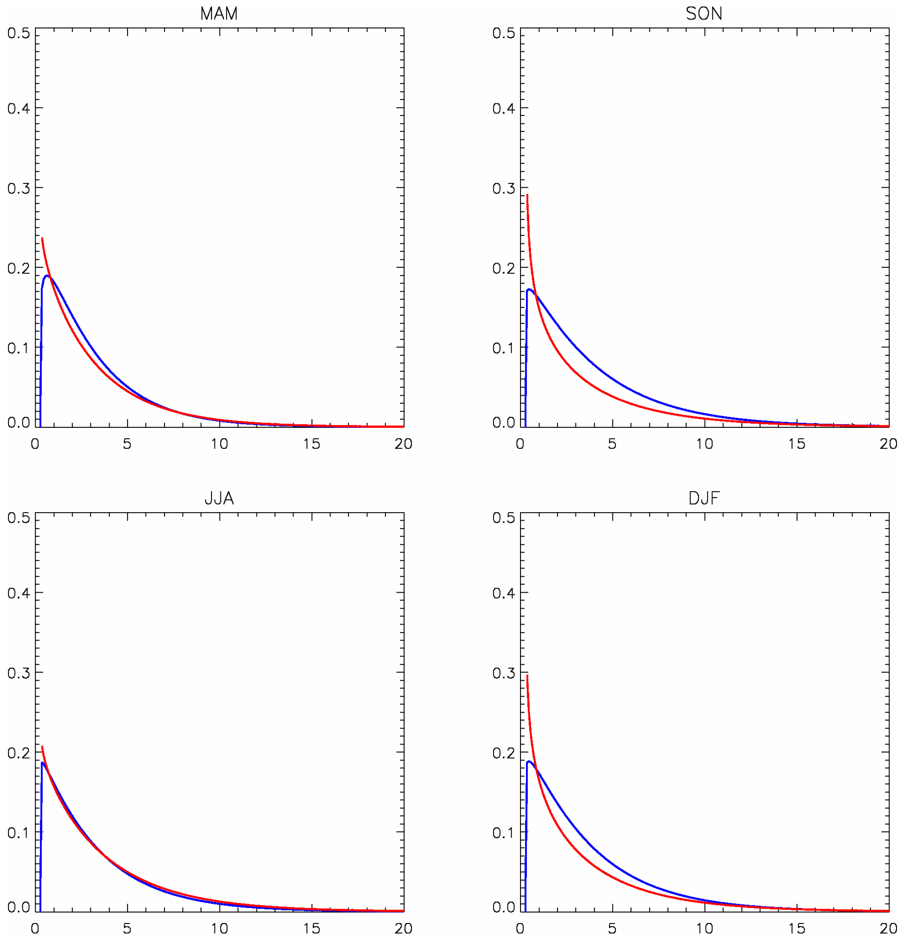

Figure 5-27: Gamma probability density functions fitted to HadCM3 SA grid-box 1 wet-day amounts, multiplied by box 1 wet-day probability.

The gamma distribution of simulated wet-day amounts, shown in Figure 5-27, is skewed substantially towards light rainfall days relative to the estimate of the observed, with almost all at less than 3mm, and anything heavier (5-15mm) occurring very rarely. This is consistent with the significantly lower gamma shape and scale parameters fitted to the simulated precipitation data compared with the observed estimates (Figure 5-26) and results in the substantial under-prediction of intense precipitation events (e.g. P95, Figure 5-26).

The difference between the 'observed' gamma parameter values that are determined using a mean series of available stations (green) are substantially different from those which are determined using the best estimate of the 'true' areal mean (blue), and lie well beyond the 95% confidence limits for all seasons except JJA. This is particularly so for the gamma shape and scale parameters, but less so for the dry-day probability and mean wet-day amount. This difference in the gamma parameters was much smaller for the UK evaluations (Figure 5-2 to Figure 5-5), and arises partly due to the smaller number of available stations but also because of the lower level of dependence between stations.

Overall, this model produces much poorer rainfall relative to the observations, in terms of amount and distribution of daily values, than it does for the UK. Although this comparison is based on a small number of boxes, it is consistent with previous evaluations that have suggested that GCMs perform better for the mid-to-high latitudes than for the tropics/sub-tropics (e.g Kharin *et al.*, 2005).

5.5.3. CCC CGCM3

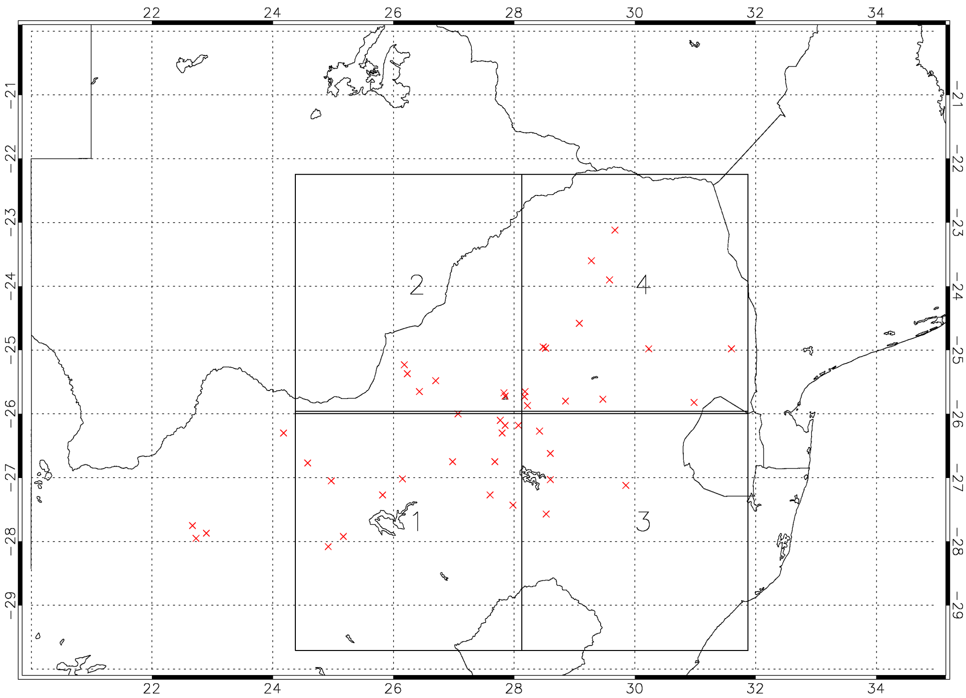

Figure 5-28: CGCM3 South African grid box positions and rain-gauge locations.

Four grid boxes covering the South African region were evaluated from CGCM3, but again, the results from only one box (grid box 3, Figure 5-28) are shown here (Figure 5-29 and Figure 5-30). None of the four grid boxes from CGCM3 for this region has a very uniform coverage of stations but box 3 is selected to be shown, despite the fact that the six stations within the box are clustered in the north-western corner. This type of station distribution is likely to occur commonly when station coverage is sparse as station coverage tends to be biased towards populated or lower-lying, and therefore more accessible, regions. It is important to note that the station distribution is as important as the overall number of available stations and the application of the methodology in a case where station coverage is biased towards one part of the grid box will result in a less robust estimate of the areal precipitation properties than if the six stations were for evenly distributed. This is because an inherent assumption of the methodology is that the characteristics of rainfall at the available stations provide a representative estimate of the characteristics of rainfall at any point in the box. This is discussed further in Section 4.4.1.2.

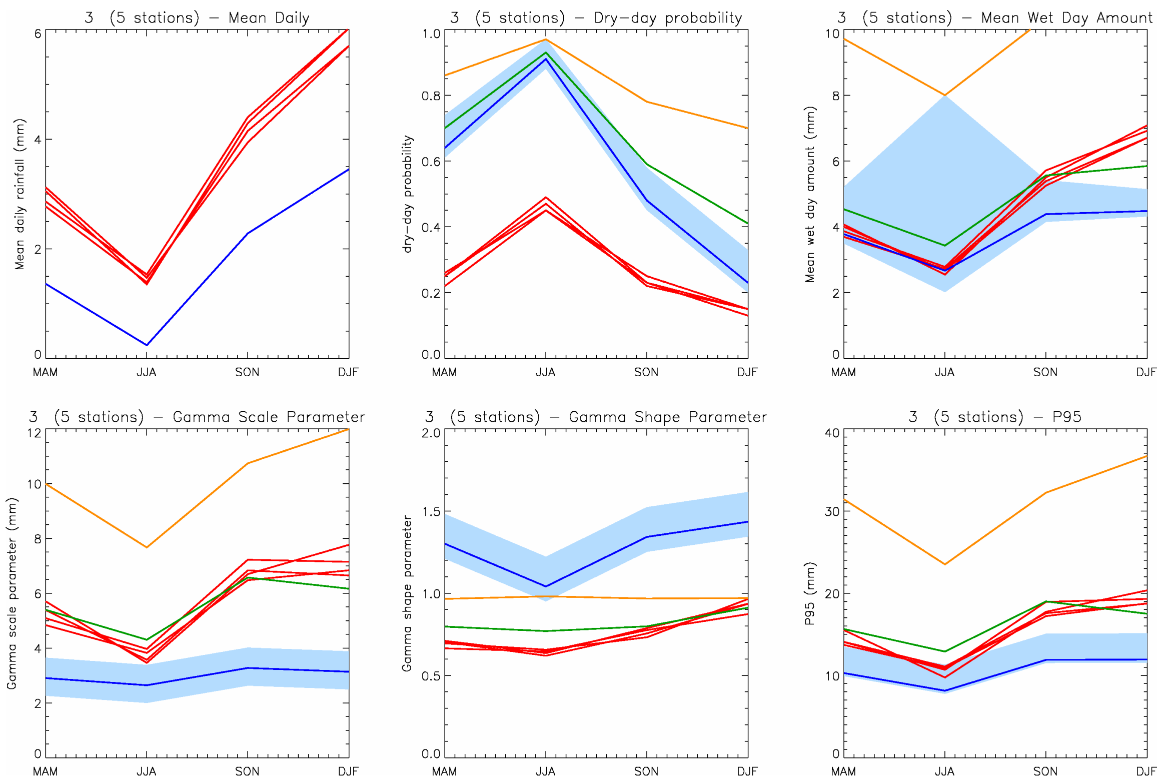

Figure 5-29: Model Evaluation Results for CGCM3 SA grid box 3. Red=Model, Blue=Observed (best estimate of true areal mean, together with 95% confidence intervals). Also shown are Yellow=average observed 'point' values, Green=observed 'areal' values as arithmetic mean of available station series'.

CGCM3 simulations for South Africa generally show a reasonably good representation of the seasonal cycle of rainfall amounts, although boxes 1 and 3 show a bias towards too much overall rainfall. There appears to be a consistent north-south variation in model performance. The northern boxes (2 and 4) show relatively good dry-day frequencies (although the seasonal cycle is slightly less pronounced than observed), whilst this is substantially underestimated for the southern boxes (1 and 3) particularly the driest seasons MAM and JJA.

The magnitude and seasonal cycle of mean wet-day amounts is relatively good for all grid boxes, although there is a tendency for this to be under-estimated in MAM. Only in Box 3 (Figure 5-29) is this over-estimated in DJF. The realism of the distribution of these wet-day values, however, varies with box and season. There is a general tendency for the scale parameter to be too high, and the shape parameter to be too low, such that too many drizzly (<1mm) daily events occur, similarly to the results from this model for the UK (see example distributions from box 3 in Figure 5-30).

For the dry-day probability, the 'best guess' estimate of true areal mean (blue) is closer to that of the modeled precipitation than that of the simple mean of available stations (green). However, for the parameters of the gamma distribution, model performance would have appeared much better if the simple areal average (green) were taken to be the observed.

Overall, CGCM3 shows a tendency for rainfall which is too frequent but too light in simulations over the UK and South Africa.

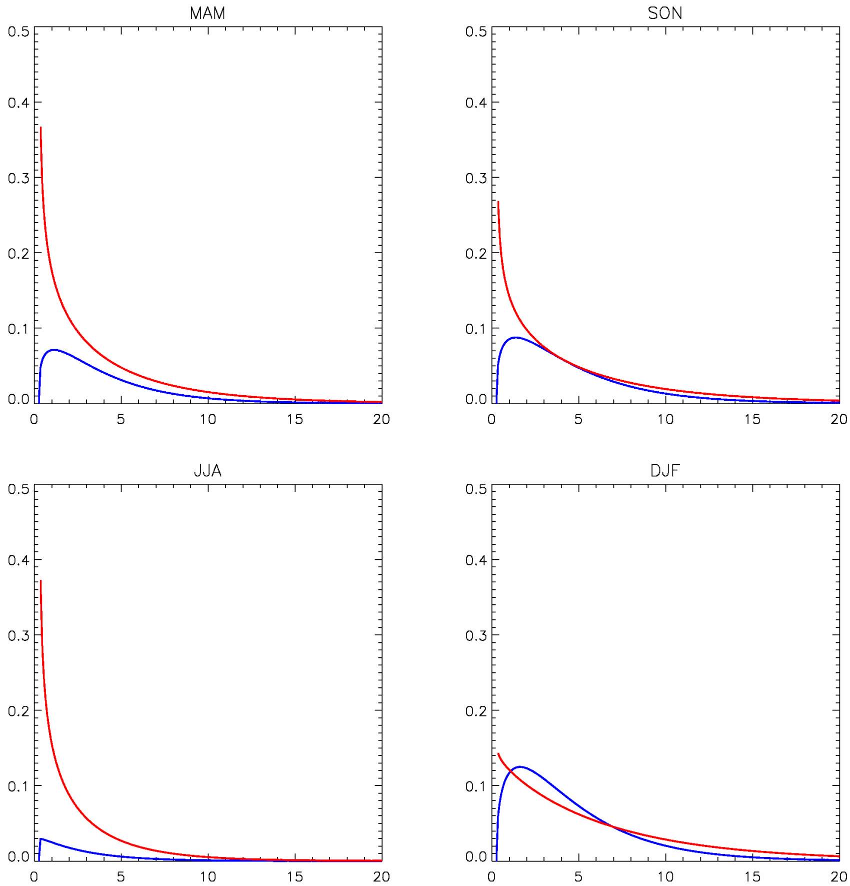

Figure 5-30: Gamma probability density functions fitted to CGCM3 SA grid-box 3 wet-day amounts, multiplied by box 3 wet-day probability.

5.5.4. NCAR PCM

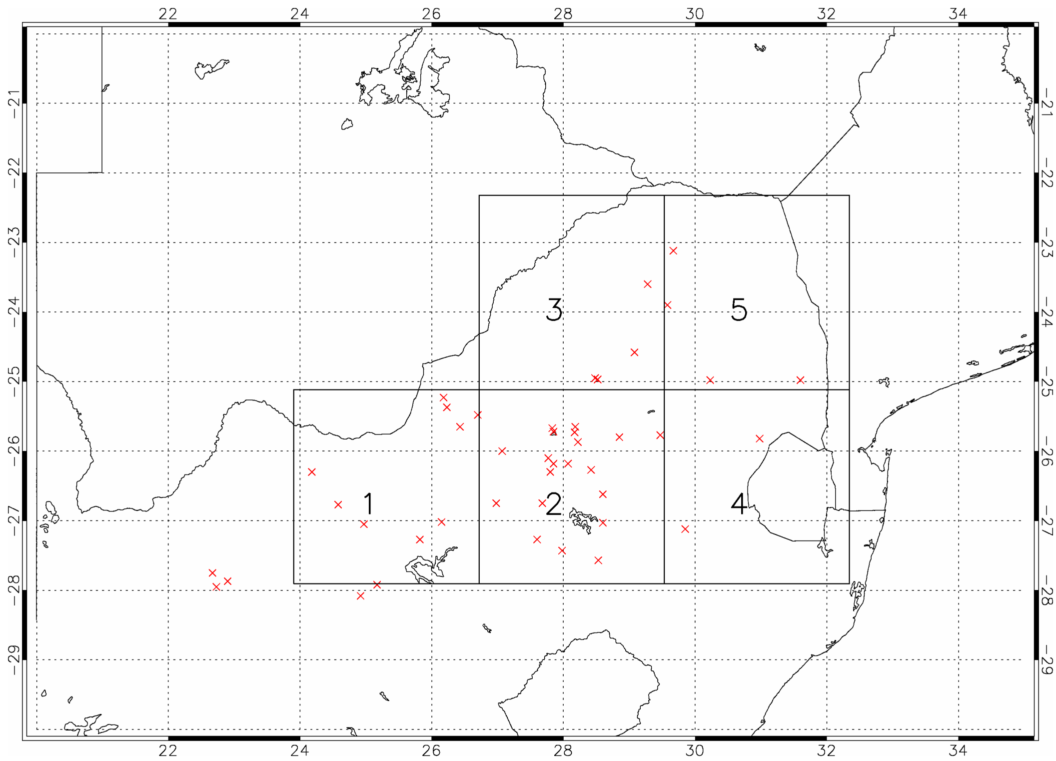

Figure 5-31: PCM South African grid box positions

The example grid box shown from the PCM evaluation for South Africa (box 4) contains just two stations (Figure 5-31). The results for this box are shown in Figure 5-32 and Figure 5-33.

The South African grid box simulations from PCM show some significant departures from the observed seasonal cycle. The observed annual cycle includes a notable 'wet' season in DJF and 'dry' season in JJA, but the model mean daily rainfall shows a moderated seasonal cycle, with the wettest season in simulations being SON, and JJA being wetter than the observed.

There is little in the way of clear or consistent errors in behaviour between seasons. JJA and SON tend to experience lower-than-observed dry-day probabilities but very realistic mean intensity. DJF and MAM do not demonstrate consistent behaviour in these parameters.

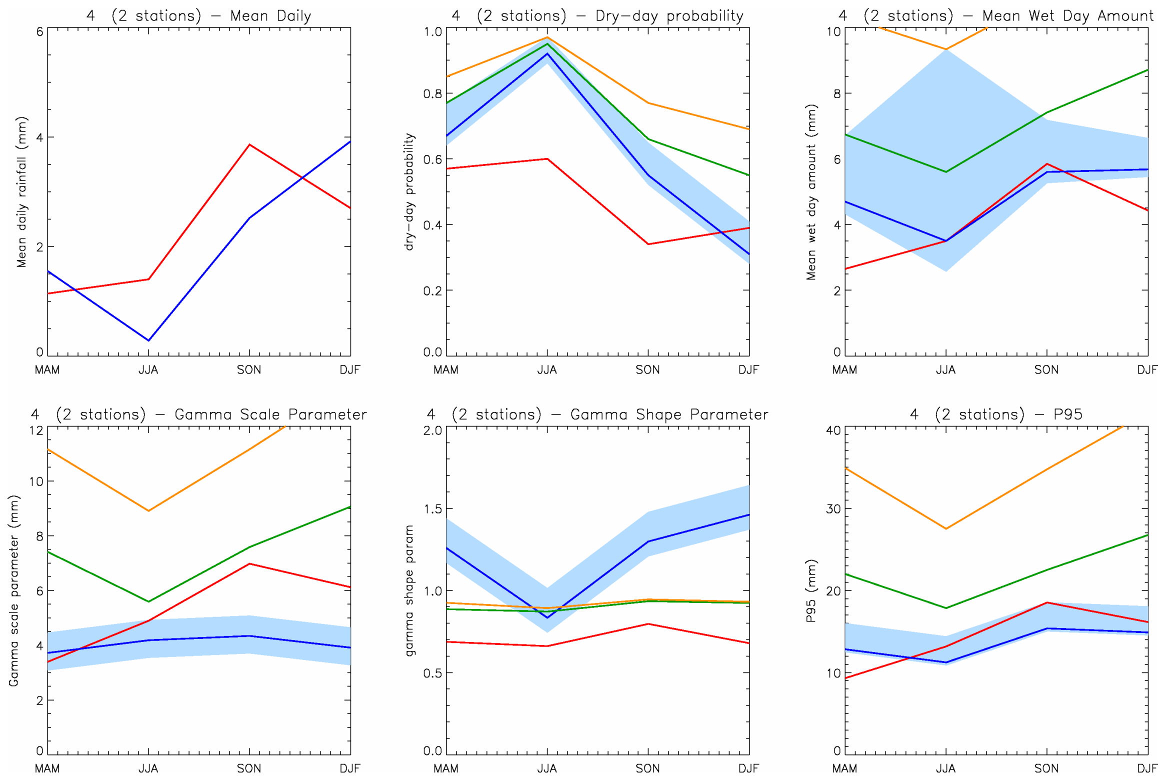

Figure 5-32: Model Evaluation Results for PCM SA grid box 4. Red=Model, Blue=Observed (best estimate of true areal mean, together with 95% confidence intervals). Also shown are Yellow=average observed 'point' values, Green=observed 'areal' values as arithmetic mean of available station series'.

As would be expected for a grid box where only 2 stations are available, the greatest differences between the different estimates of observed areal rainfall characteristics (green and blue lines) are seen here, particularly for the mean wet day amount. Of course, this also means that the additional uncertainties to those which are shown on the 'best guess' estimates are the largest.

In terms of the gamma distribution parameters of wet-day amounts, JJA is the most realistic. Whilst the number of wet-days is over-estimated, the proportion of days of different rainfall intensities is relatively realistic. The other seasons tend to experience a shape parameter which is too low, consistent with the tendency for too high a proportion of 'drizzly' days, but also a scale parameter which is too high in all seasons except MAM. This means that the distribution, when scaled according to the wet-day probability, tends to be too 'concave', with higher proportions of light and very heavy events, but less frequent moderate events (Figure 5-33).

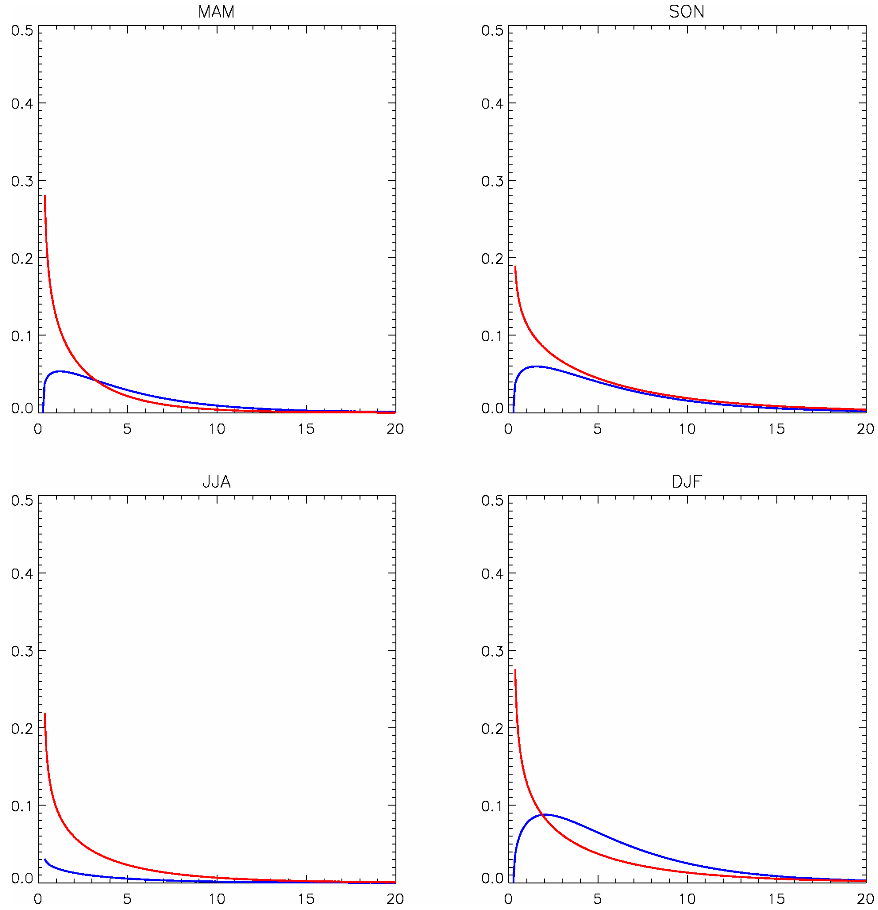

Figure 5-33: Gamma probability density functions fitted to PCM SA grid-box 4 wet-day amounts, multiplied by box 4 wet-day probability.

5.6. Discussion and Summary

The methods developed for estimating various characteristics of observed, areal rainfall using limited available station data have allowed a more explicit evaluation of the characteristics of daily rainfall variability from climate model simulations than has previously been possible. These evaluations are now discussed in terms of the information they yield about the models, and then in terms of the methods applied.

5.6.1. Discussion of Model Performance

Previous model evaluations (Mearns *et al.*, 1995; Chen *et al.*, 1996; Dai *et al.*, 1999, Sun *et al.*, 2006) have indicated that a common problem with simulations is that precipitation occurs too frequently but at too low an intensity; that is, too little and too often (Sun *et al.*, 2006).

This study has been limited to two relatively small study regions; the UK and northern South Africa, which is not a sufficient proportion of the world to draw general conclusions about the performance of the three models considered here. The results are discussed in the context of earlier evaluations, and in terms of the additional detail that they offer.

The more explicit evaluation of the frequency and distributions of rain days in this study has demonstrated that two of the three models (PCM and CGCM3) demonstrate the commonly observed tendency towards 'too little, too often' rainfall, with both dry-day frequency and mean intensity appearing to be low compared to the observed. Further to this general behaviour, the comparison of gamma distribution parameters of wet day amounts has highlighted some important characteristics of the distribution of modeled rainfall. CGCM3 has a strong tendency towards 'drizzle', (days receiving 3mm of rainfall or less), the common problem seen by many earlier generation and new generation climate models. This causes the mean wet-day amount (or mean intensity) and the dry-day frequency to be too low, whilst mean rainfall, overall, can still be realistic. This behaviour has been recognized in several other evaluations of new and old versions of the CGCM model (Osborn and Hulme, 1998; Dai and Trenberth, 2004; Dai *et al*, 2006).

PCM also has a tendency towards too few dry-days, and low mean wet-day amount (mean intensity), but the distribution of values indicates that this is not due to 'drizzle' but to a more general inclination towards too many wet days across the full range of intensities

HadCM3 shows less systematic bias. The rainfall is of reasonable magnitude and frequency, and the distribution of values is generally very good compared to PCM and CGCM3. This is consistent with the results of Osborn and Hulme (1998), who found that HadCM3 performed relatively well with respect to the characteristics of daily rainfall compared to 11 other models. However, the seasonal cycle in these HadCM3 grid boxes are much less realistic, as the errors (relatively small in the UK and larger in South Africa) are not consistent between seasons.

Comparatively, the model performance for all models over South Africa is relatively poor. The distributions of wet-day amounts that were well-reproduced for the UK become much less skillful for South Africa, and are too 'drizzly'. CGCM3 makes the best reproduction of the seasonal cycle of total rainfall, but still struggles to distribute this rainfall realistically in time. Previous studies such as Kharin *et al.* (2005) have found that climate models reproduce the characteristics of rainfall in the tropics and sub-tropics with less skill than in the mid-to-high latitudes. This is because tropical rainfall is more convective in its nature, and its reliable modeling therefore relies on realistic representation of the complex convection process. Convection is particularly difficult to model reliably at the limited resolution of climate models and results have been demonstrated to be highly sensitive to the parameterisations of convection that are used (Scinocca and McFarlane, 2004; Kharin *et al.*, 2005).

5.6.2. Discussion of Model Evaluation Method

The novel methods presented and applied in this thesis have allowed quantitative evaluations of the variability of daily rainfall to be made, by estimating the parameters of the 'true' areal mean observed series.

For UK, where station density is relatively high, the parameters of the *n*-station series (green) generally lie within, or near, the confidence limits of the estimated 'true-areal mean', such that the process of estimation does not add a great deal of improved accuracy to the evaluation.

In the case of South Africa, however, where the station density is lower and in some cases unevenly distributed, the difference between the mean of the available stations and the estimated mean of the area is more marked. In these cases, the benefits of the methodology developed here are clear. The results of the CGCM3 evaluation for South Africa, for example, might have suggested that the model accurately represents the distribution of rainfall if a mean series of available stations was used as the 'observed' record. When compared to the distribution of values estimated for the true areal mean, however, it is clearer that the rainfall simulated by the models is not so realistic.

Whilst the benefits of the method increase for regions which are less densely gauged, the uncertainties associated with the best estimates of 'true' areal mean are also increased. These additional uncertainties are difficult to quantify, as they vary considerably from one situation to another, but are discussed in more detail in Chapters 3 (Section 3.5.2) and 4 (Section 4.4.3). In this particular example, the estimates for JJA rainfall are particularly uncertain because the frequency and amount of rainfall is so low as to make it difficult to reliably fit a gamma distribution. It is recommended that the method is applied to regions or seasons with a daily average of less than 0.3mm with particular care.

Greater uncertainties in observations for areas that are sparsely gauged are, however, inevitable as the fewer data that are available for a region, the larger the proportion of the information that has to be inferred. It is hoped, therefore, that whilst relatively large uncertainties are associated with these estimates, these still represent a better estimate of areal precipitation variability than the areal averages that rely only on a limited number of stations. The uncertainties should be kept in mind when these approaches are applied, even when they are more qualitative than quantitative.