South Eastern Australian Climate initiative

Final report for Project 1.3.5

Further Development of Statistical Downscaling Methodology

Principal Investigator: Steve Charles
CSIRO Land and Water, Private Bag 5, Wembley WA 6913
Steve.Charles@csiro.au
Tel: 08 9333 6795
Fax: 08 9333 6499

Co-Authors: Guobin Fu
CSIRO Land and Water, Private Bag 5, Wembley WA 6913
Guobin.Fu@csiro.au

Completed: January 2008

Abstract
The stochastic downscaling framework implemented in this, and related, projects simulates rainfall as a two–stage process, for rainfall occurrences and amounts respectively. The Nonhomogeneous Hidden Markov Model (NHMM) simulates multi-site daily rainfall occurrences. Then a regression approach is used to model a station's empirical rainfall distribution conditional on neighbouring station rainfall occurrences. We assess an extended amounts model that includes atmospheric predictor covariates in the regression model, in addition to neighbouring station rainfall occurrences. We also add resampling from a gamma distribution, rather than the empirical distribution. Results indicate that the extended amounts model versions often better reproduce observed validation period rainfall statistics.

Another limitation of the current NHMM is a calibration restriction to 30 stations or less. A version of the NHMM developed separately by collaborators at the International Research Institute for Climate and Society (IRI) does not have this restriction, allowing calibrations to networks with a much larger numbers of stations. Here we show the IRI-NHMM can reproduce validation period rainfall statistics for a network of 132 stations.

Significant research highlights, breakthroughs and snapshots

- Extended versions of the rainfall amounts component of the stochastic downscaling model, using atmospheric predictors in addition to neighbouring station rainfall occurrence, and a gamma distribution instead of the empirical distribution, are often able to better reproduce the daily rainfall distribution of the validation period. These extended rainfall amounts model versions will be used in the generation of projected rainfall series and the differences assessed in Project 2.1.4.
- The IRI-NHMM provides the ability to downscale to a much larger network of stations, although the algorithms it uses for reproducing between-site spatial correlations are not as advanced as those of the CSIRO-NHMM. Testing on a network of 132 stations produced adequate results.

Statement of results, their interpretation, and practical significance against each objective

> **Objective 1:** Determine which, if any, deficiencies in NHMM performance require further model development.

Significant changes can occur in individual station daily rainfall probability density functions (PDFs) over time. Thus a key measure of the performance of any statistical downscaling approach is its ability to reproduce the rainfall statistics of a period not used in calibration. Here we use the 1958-1985 period to assess the NHMMs that were calibrated for 1986-2005 in Project 1.3.4 *Calibration of Statistical Downscaling Models*. Six variants of methods for conditioning multi-site daily rainfall amounts on neighbouring station rainfall occurrences and atmospheric predictors, resampling from gamma or empirical distributions, were assessed to determine how well such changes are captured.

The original CSIRO-NHMM used only neighbouring station rainfall occurrences and resamples station amounts from observed empirical distributions (Charles *et al.* 1999).

In addition, here we include two modifications: (i) conditioning on atmospheric predictors and (ii) sampling from a gamma distribution fitted to observed station daily rainfall (by weather state). Thus the six variants involve conditioning on:

- (i) neighbouring station rainfall occurrences only with empirical distribution
- (ii) neighbouring station rainfall occurrences only with gamma distribution
- (iii) neighbouring station rainfall occurrences and atmospheric predictors with empirical distribution
- (iv) neighbouring station rainfall occurrences and atmospheric predictors with gamma distribution
- (v) atmospheric predictors only with empirical distribution
- (vi) atmospheric predictors only with gamma distribution

Note the predictor sets used for conditioning these amounts models are independent of those used for NHMM calibration. The candidate set of predictors for amounts conditioning was restricted to specific humidity and dew point temperature depression at 850, 700 and 500 hPa. A stepwise multiple linear regression procedure was applied to select the conditioning predictor from these candidates, for each station and for each weather state.

Figures 1 to 5 present, for five example stations, quantile-quantile plots of observed versus downscaled validation period daily rainfall. The versions using the fitted gamma distributions, rather than the empirical distributions, often better reproduce the upper tails of the distributions (i.e. the very high daily rainfalls that occur rarely). The gamma distribution can also sometimes improve the overall distribution, but more often the addition of the atmospheric predictors to the covariates used, in combination with the gamma distribution, improves performance the most.

> **Objective 2:** Assess the CSIRO NHMM against a NHMM version produced by collaborators at the International Research Institute for Climate and Society (IRI), New York, USA.

We investigate the IRI-NHMM performance for a 132 station network in terms of reproduction of probability of rainfall occurrence (i.e. wet-day frequencies) and log-odds ratio. The log-odds ratio is similar to a measure of correlation but measures association in binary (i.e. wet or dry) data. We compare the winter weather states of the CSIRO-NHMM (Project 1.3.4 report Figure 4) against that of the IRI-NHMM in Figure 6. The rainfall patterns of the 132 station model correspond to the states seen in the original 30 station model, as do the resultant composite predictor fields.

Figure 7 shows that the 132 station IRI-NHMM can reproduce the winter rainfall probabilities of the fitting period with a slight bias for the validation period. A similar bias was seen in the original 30 station CSIRO-NHMM (Project 1.3.4 report Figure 8). Similarly, log-odds ratios for the 132 stations are adequately reproduced although the non-spatial IRI-NHMM parameterisation leads to underestimation for station pairs with high correlation (Figure 8, compare to Project 1.3.4 report Figure 9). Overall the IRI-NHMM can perform reasonably well for a 132 station network. More detailed investigation of inter-site rainfall statistics would be needed if particular applications where high spatial consistency is crucial, such as sub-catchment hydrological modelling, were envisaged.

Summary of methods and modifications (with reasons)

- Six versions of the daily multi-site amounts model were assessed and the version using additional atmospheric covariates as well as resampling from a gamma distribution was often found to perform the best.
- The NHMM version developed by the IRI has the advantage that it is not limited to 30 rainfall stations, as is the CSIRO NHMM code. However its ability to model the between-site spatial characteristics is parameterised in a less rigorous way and hence such characteristics may not be adequately reproduced for all applications. Here the IRI-NHMM calibrated to the same 30-station network, and also to a 64 and 132 station network (necessarily accepting stations of lesser data quality) performed well. Only results from the 132 station version are presented here.

Summary of links to other projects

- The statistical downscaling model improvements produced in 1.3.5 are essential for providing the best NHMM that will be applied to:
    - understanding current regional hydroclimate (Projects 1.4.2, 1.4.3, 1.5.2, 1.5.3)
    - climate change projection at the regional scale (Projects 2.1.3, 2.1.4).
- The improved NHMM will be compared to the models developed in Projects 1.3.1 and 1.5.5.

Publications arising from this project

Kirshner, S. 2007. Learning with Tree-Averaged Densities and Distributions (conference paper, uses SEACI dataset and acknowledges S.Charles)

Acknowledgement

The IRI version of the NHMM used in the project was developed by Dr Sergey Kirshner, University of Alberta, Canada in collaboration with Dr Andrew Robertson, IRI for Climate and Society, New York, USA. Their on-going assistance is gratefully acknowledged. This project is funded by the South Eastern Australian Climate Initiative.

Recommendations for changes to work plan from your original table

None.

References

Charles SP, Bates BC, Hughes JP. 1999. A spatio-temporal model for downscaling precipitation occurrence and amounts. *Journal of Geophysical Research – Atmospheres* **104(D24):** 31657-31669.

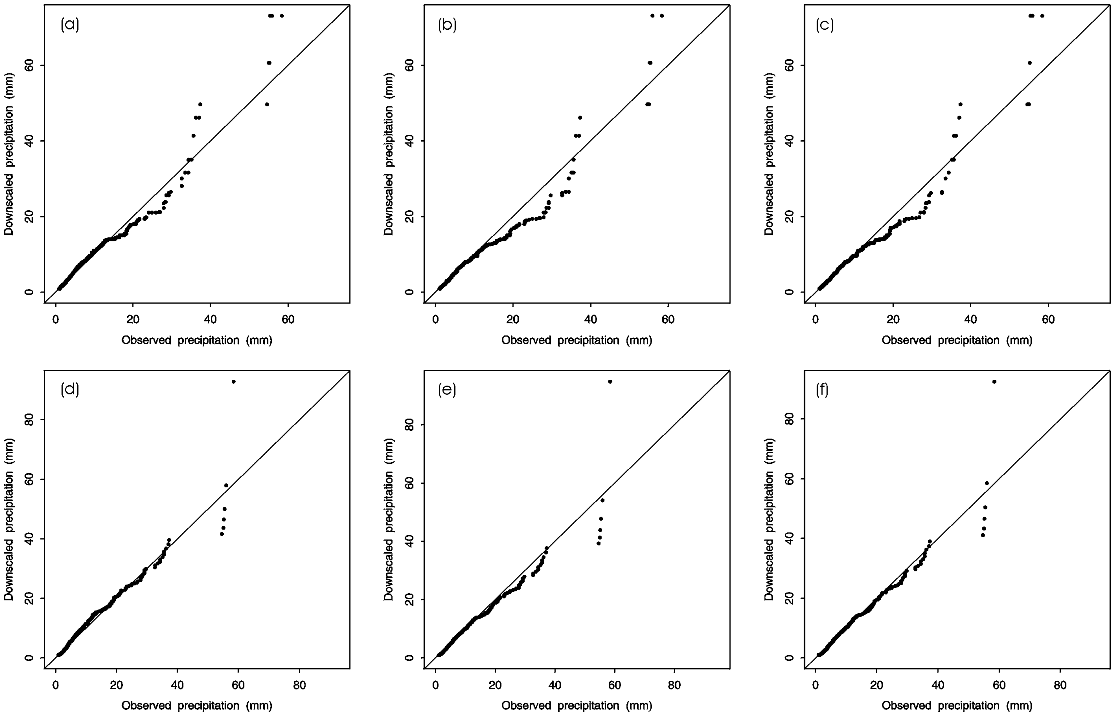

Figure 1. Station 49048 daily amounts quantile-quantile plots for amounts model versions using: (a) neighbouring station rainfall occurrences only with empirical distribution; (b) atmospheric predictors only with empirical distribution; (c) neighbouring station rainfall occurrences and atmospheric predictors with empirical distribution; (d) neighbouring station rainfall occurrences only with gamma distribution; (e) atmospheric predictors only with gamma distribution; and (f) neighbouring station rainfall occurrences and atmospheric predictors with gamma distribution.

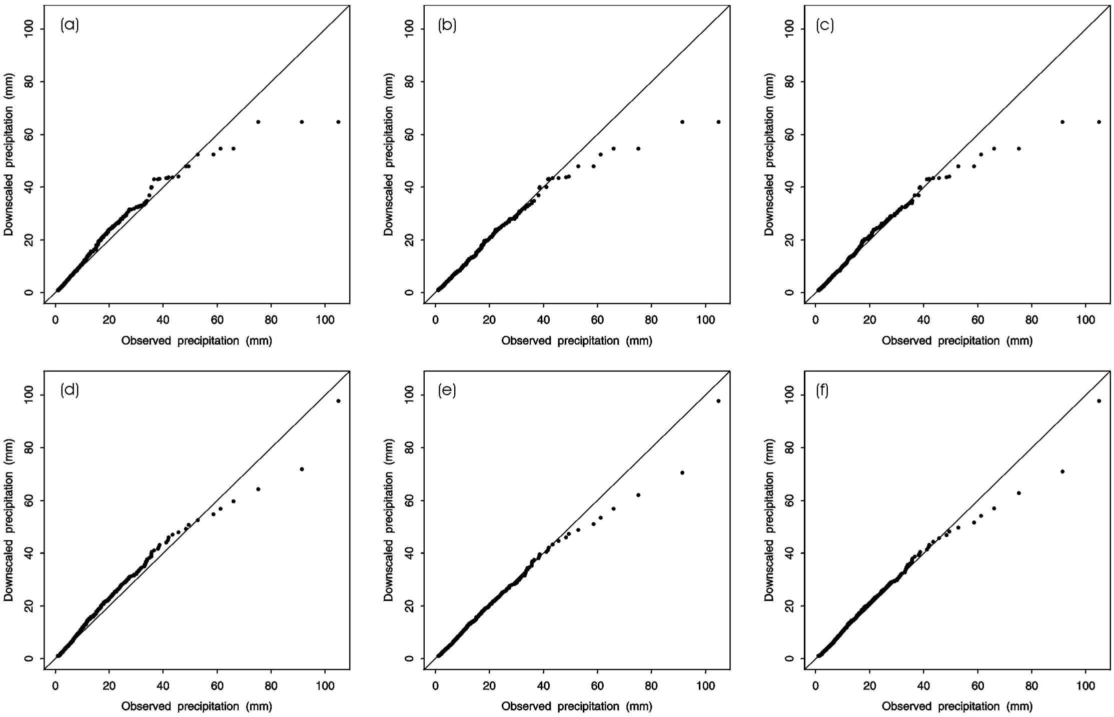

Figure 2. As Figure 1, for Station 70014.

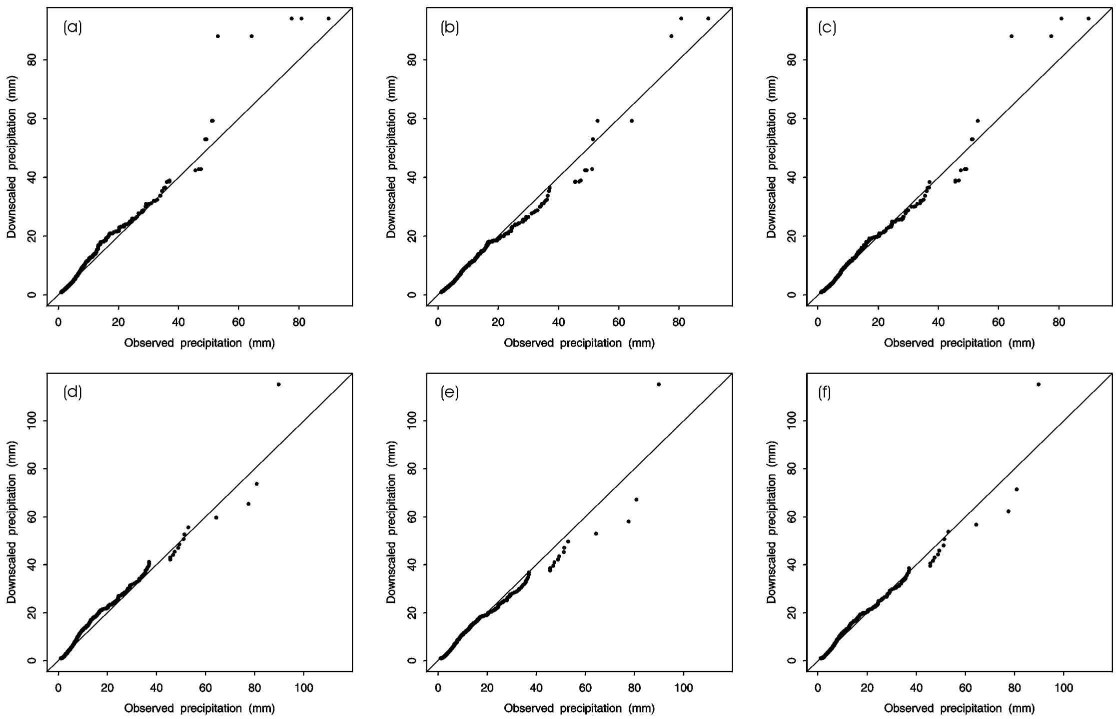

Figure 3. As Figure 1, for Station 70054.

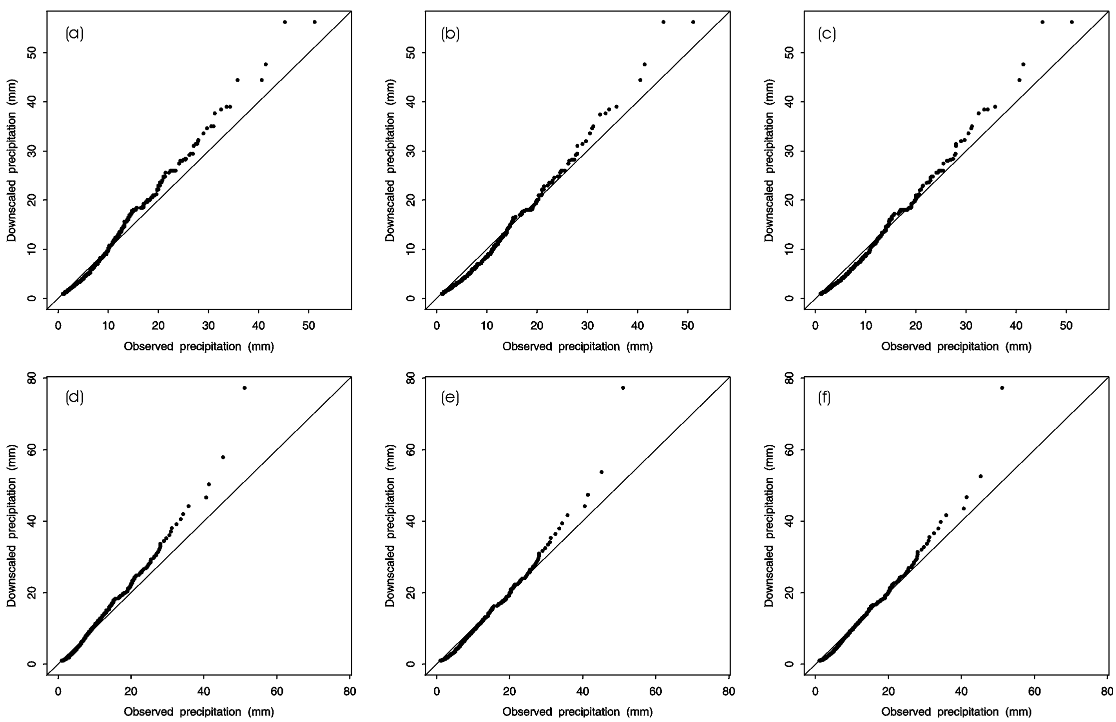

Figure 4. As Figure 1, for Station 74087.

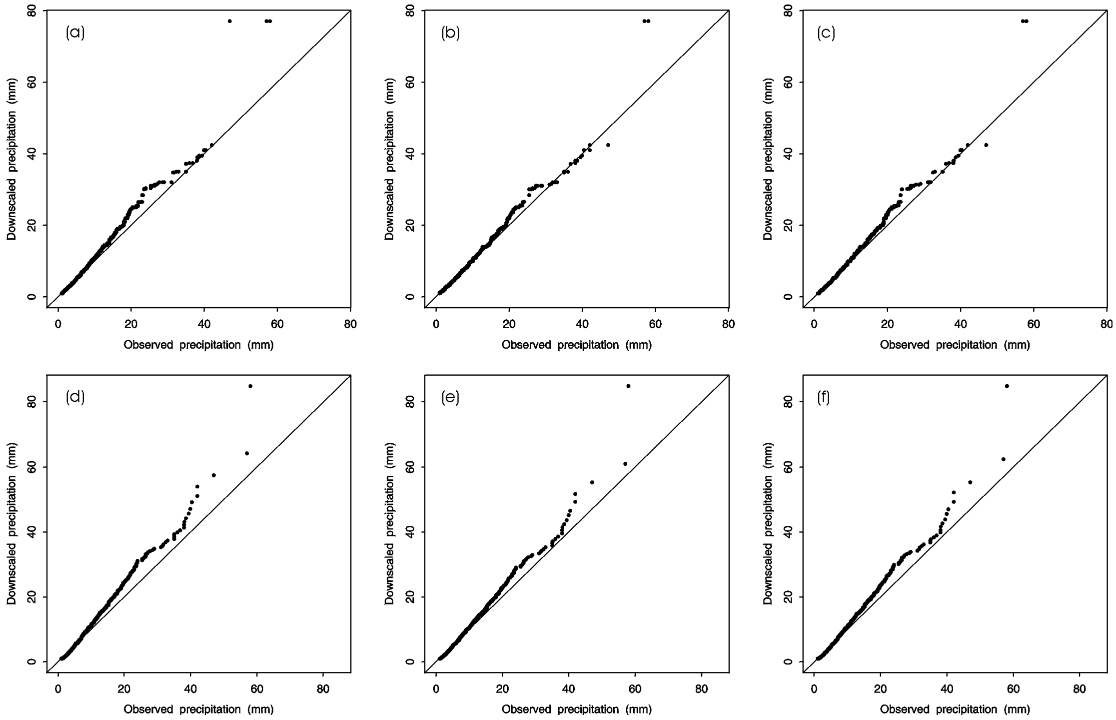

Figure 5. As Figure 1, for Station 88011.

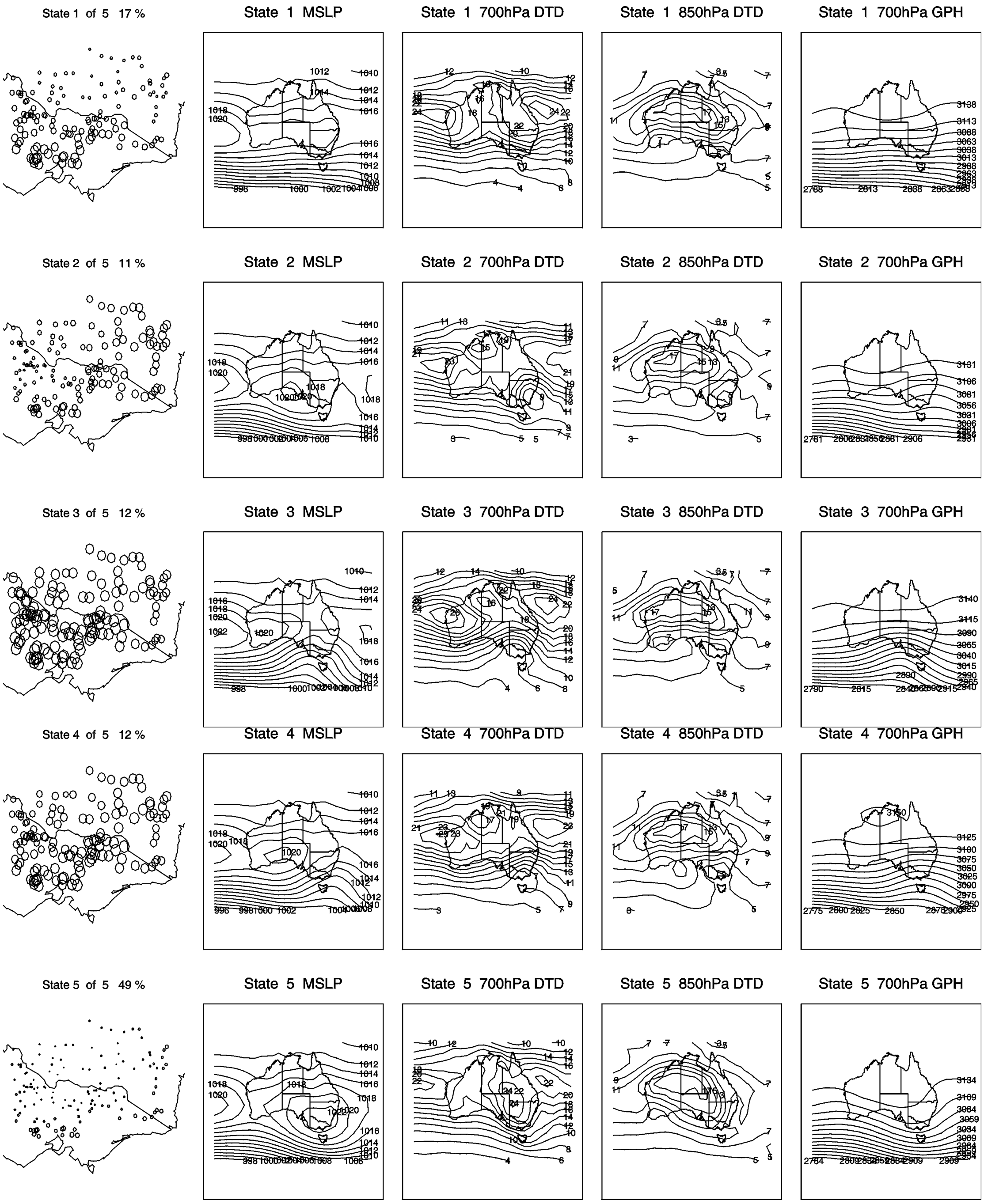

Figure 6. Winter weather states of 132 station IRI-NHMM.

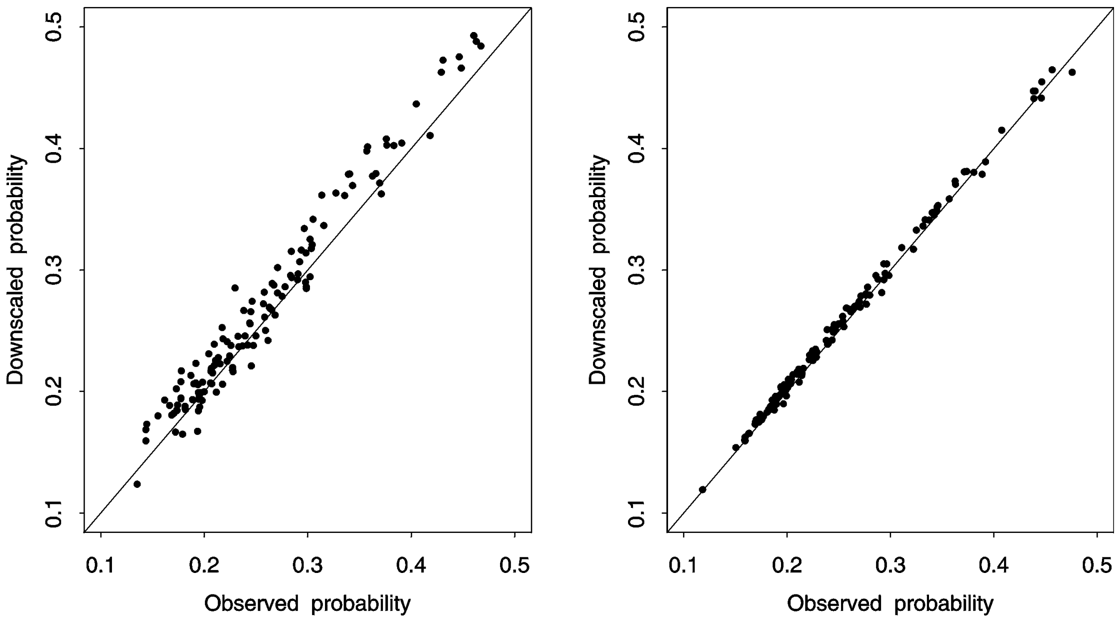

Figure 7. Probability 132 stations (a) validation period (b) fitting period.

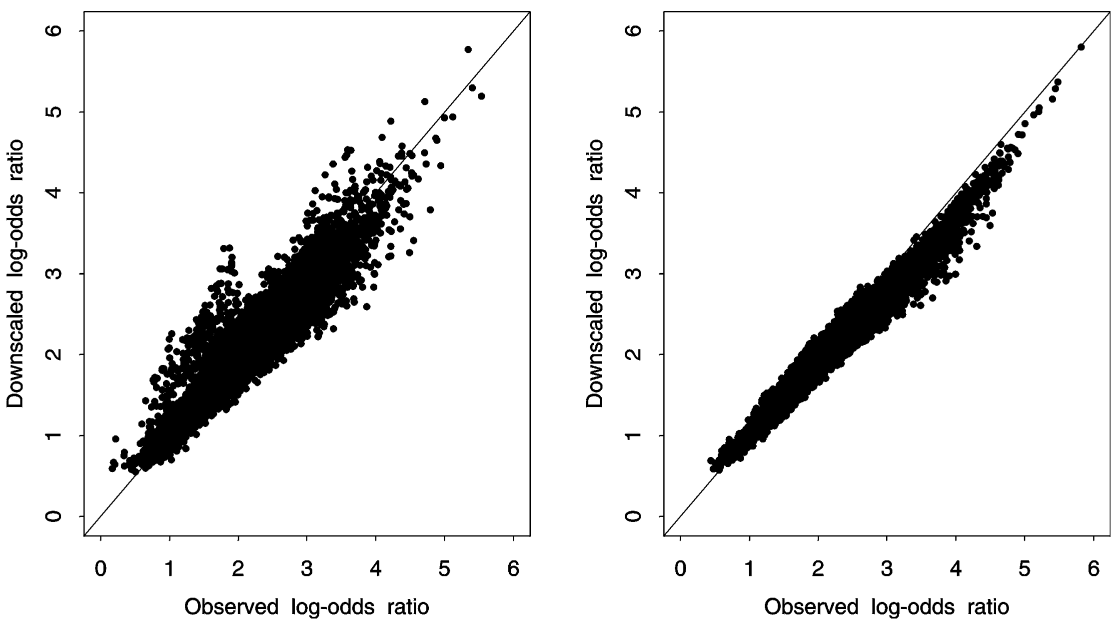

Figure 8. Log-odds 132 stations (a) validation period (b) fitting period.

Project Milestone Reporting Table

To be completed prior to commencing the project				Completed at each Milestone date	
Milestone description[1] (brief) (up to 33% of project activity)	Performance indicators[2] (1- 3 dot points)	Completion date[3] xx/xx/xxxx	Budget[4] for Milestone ($) (SEACI contribution)	Progress[5] (1- 3 dot points)	Recommended changes to workplan[6] (1- 3 dot points)
1. Assess deficiencies in CSIRO NHMM performance	Properties of rainfall occurrence and amounts requiring improvement identified.	1/8/2007	2.5	Completed.	
2. Modify CSIRO NHMM to address deficiencies	Modified NHMM calibrated and assessed.	31/12/2007	7.5	Completed.	
3. Calibrate IRI NHMMs	Calibration of IRI NHMMs performed.	1/11/2007	5	Completed.	
4. Compare CSIRO NHMM performance to IRI NHMM performance	Relative performance of NHMMs assessed. Written report (4-6 pages).	31/12/2007	5	Completed (this report is the written report)	