On the Horizontal Scale of Topographic Dependence of Monthly Australian Precipitation

J.J. Sharples and M.F. Hutchinson

Centre for Resource and Environmental Studies, Australian National University.

Email : jasons@cres.anu.edu.au

Abstract: Consideration of the scale of interaction between precipitation and topography is important for the accurate interpolation of rainfall in mountainous areas and also provides insight into the physical processes involved. In this paper we use trivariate thin-plate smoothing splines to interpolate monthly rainfall over two subregions of the Australian continent, incorporating different overriding climatic conditions and rainfall types. The interpolations are based upon elevations derived from elevation grids of various resolutions. All the grids are local averages of version 2.0 of the 9 second resolution digital elevation model of Australia. The results suggest that the optimal scale of interaction between precipitation and topography, as it pertains to the interpolation of precipitation in Australia, is between 5 and 10 kilometres. This is in agreement with results of similar studies dealing with daily precipitation.

Keywords: *Precipitation; Topography; Interpolation; Thin-Plate Smoothing Spline.*

1. INTRODUCTION

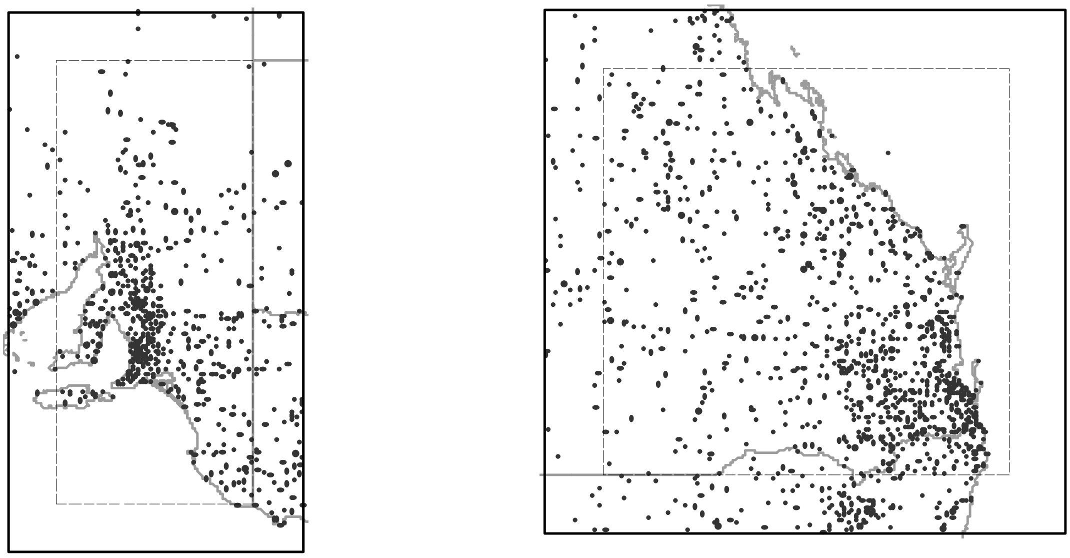
Precipitation exhibits complex spatial behaviour and it is well known that this behaviour is often related to the underlying topography. Determining the scales of the interaction of observed precipitation with topography thus has bearing on the spatio-temporal analysis of precipitation and can provide insight into the nature of the precipitation processes involved. In particular, knowledge of the scale of dependence of precipitation on topography can be used to optimise interpolation methods that incorporate topographic dependence for estimating the spatial distribution of quantities related to precipitation. Such interpolatory methods play a pivotal role in a wide range of applications concerned with assessing the impacts of climate on agriculture, ecology, hydrology and tourism (Hutchinson [1995a], Hulme, *et al*. [1995], Houghton *et al.* [1996], Ruddell, *et al*. [1990]). In addition, since physically based precipitation models present forecasts in the form of gridded surfaces that are typically made at grid resolutions from tens to hundreds of kilometres, methods for interpolating precipitation observed at discrete locations has an important part to play in the calibration and validation of such models.

In this paper we specifically address the problem of determining the optimal degree of horizontal resolution suitable for the interpolation of monthly point precipitation values. Given the physical characteristics of the processes involved during a precipitation event it is intuitively clear that the layers of air between the rain bearing clouds and the underlying topography act as a diffusive buffer. Thus one would expect the precipitation patterns to conform to a broader scale representation of the underlying topography. Determination of the appropriate degree of horizontal scaling for the interaction between monthly precipitation totals and topography in Australia is hence the main focus of this paper.

2. STUDY REGIONS AND DATA

We present results for two subregions of the Australian continent. The first region of interest is the portion of South Australia/Victoria falling within the limits of 137^o and 141^o longitude and -38^o and -29^o latitude. The second region is the portion of Queensland/New South Wales falling within the limits of 147^o and 154^o longitude and -29^o and -22^o latitude. Margins of 1^o were added to these limits to reduce errors due to edge effects.

The two regions with margins are shown in Figure 1 along with the precipitation data locations. We will use the abbreviations SA to refer to the first region and SQ to refer to the second region. These regions were chosen since they both contain significant topography. The climate of the SQ region, however, is subtropical/tropical, while the climate of the SA region is temperate. The precipitation in the SA region usually results from frontal activity while in the SQ region the precipitation can result from both frontal activity moving up the east coast as well as seasonal convective activity originating in the tropics. The precipitation data used in this



(a) (b)

Figure 1. Locations of data points for (a) SA region and (b) SQ region. The dashed lines frame the areas of study. The solid lines enclose those data points within the study areas plus margins.

study were obtained from the Bureau of Meteorology and consisted of monthly totals for the year 2000. For the SQ region the data consisted of 1023 points within the margins, 830 of which were within the limits defined above. Of these 830 points, 60 were randomly removed to provide a validation data set for the fitted spline. For the SA region the data consisted of 566 points within the margins, 449 of which were within the specified limits. Of these 449 points, 40 were randomly removed for validation purposes.

The square root transformation was applied to the precipitation values before the interpolation to remove the natural skewness in the distribution of rainfall values. Application of this transformation also helped facilitate the identification of bad data values.

The monthly precipitation data was interpolated using the ANUSPLIN 4.3 package, which is a collection of FORTRAN routines for calculating thin-plate smoothing splines (Hutchinson [2003]). A new feature of the ANUSPLIN 4.3 package is a facility for the systematic identification and removal of suspected bad data points. Data values that differ from the fitted surface by more than 3.6 standard deviations are flagged. These data values can then be compared to values at nearby locations. In this way erroneous data can be identified and removed from the data set. Typically, one often finds spurious zeros that have been confused with a missed reading. Given this new feature, the full data sets (before removing validation points) were subjected to initial spline analyses using elevation data derived from the 9 second DEM. These initial analyses then provided a set of suspect data points for each region that was excluded from all of the subsequent spline analyses. It should be noted that many of the suspect zero data points would not have been identified without the aid of the square root transformation.

The elevation data were derived from local averages of version 2.0 of the 9 second (approx. 250m) resolution digital elevation model of Australia (Hutchinson et al. [2001]). The local averages of the DEM were obtained using the GRDGEN routine (Hutchinson, [2002]), which averages the DEM over grid cells of a specified size. Thirteen different grid resolutions were used, ranging from 250m to 90km. Scale specific elevation data were then attributed to each data location using the INTGRD routine (Hutchinson, [2002]), which applies biquadratic spline interpolation to each of the locally averaged DEMs. In addition, the station heights as recorded in the Bureau of Meteorology's database were also used to test how ultra fine resolution elevation data effects the interpolation of precipitation. In this way fourteen sets of precipitation/elevation data were obtained for the regions, each pertaining to a different horizontal topographic scale.

The spline interpolation for the SA region was based on 400 knots chosen by successively removing closest data points. Similarly, for the SQ region, the surface fitting was based upon 900 knots. The use of knots, rather the full data set, as a basis for spline interpolation, bypasses the effects due to short-range correlation within the data and can thus lead to more stable surfaces. The ANUSPLIN 4.3 package allows for the automatic selection of a specified number of knots using the SELNOT routine.

3. SMOOTHING SPLINE METHOD

The measured precipitation values are assumed to be realisations of the model

$$r_i^{1/2} = f(x_i, y_i, h_i) + \varepsilon_i$$

where r_i denotes the precipitation recorded at the location whose longitude and latitude are x_i and y_i respectively, and whose elevation is h_i. The ε_i represent random error terms, assumed to satisfy

$$\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)^T = N(0, \sigma^2 I).$$

Units of degrees were used for the latitude and longitude of the data locations, while units of kilometres were used for the elevations. This is in keeping with the commonly accepted atmospheric horizontal and vertical distance scales of 1000km and 10km respectively (Daley [1991]) and with the results of other studies concerned with the effect of the vertical exaggeration of elevation on the interpolation of precipitation (Hutchinson and Bischof [1983], Hutchinson [1995b]).

The general thin-plate smoothing spline estimate of the function f is obtained by minimising

$$1/n \, \| \mathbf{r}^{1/2} - f(\mathbf{x}, \mathbf{y}, \mathbf{h}) \|^2 + \lambda J_m(f)$$

over a suitable class of functions. The first term in the above expression is the average squared Euclidean distance between the observed data and fitted values, and $J_m(f)$ is the m^{th} order roughness penalty consisting of the integral of squared m^{th} order partial derivatives of f. In this study we will take m=2 so that we are considering smoothness in terms of second order partial derivatives of the function f. The parameter λ determines a balance between the fidelity to the data and the degree of smoothness of the fitted spline function f. This parameter is usually determined automatically by minimising the generalised cross validation (GCV). The GCV provides a reliable measure of the predictive error of the fitted surface that is calculated from the data by implicitly withholding each data point in turn from the fitting procedure. For further details on thin-plate smoothing splines the reader is referred to Wahba [1990] and Hutchinson [1995b].

The SPLINB routine in the ANUSPLIN 4.3 package facilitates the calculation of minimum GCV thin-plate smoothing splines based on knot sets.

4. RESULTS

The square root of the GCVs and the residual mean square errors for the square roots of the precipitation values in the validation data sets were obtained as outputs from the SPLINB routine.

The effect that the horizontal resolution of topography has on the interpolation of rainfall in the SA and SQ regions can be seen in figure 2 and figure 3, respectively. These figures show plots of the GCV versus horizontal resolution for each of the months of the year 2000. The various GCV curves have been vertically translated to elucidate their comparison. The vertical axes have been left unmarked for this reason. In doing the translations, however, we have preserved the relative positions of the monthly curves with respect to the vertical axis.

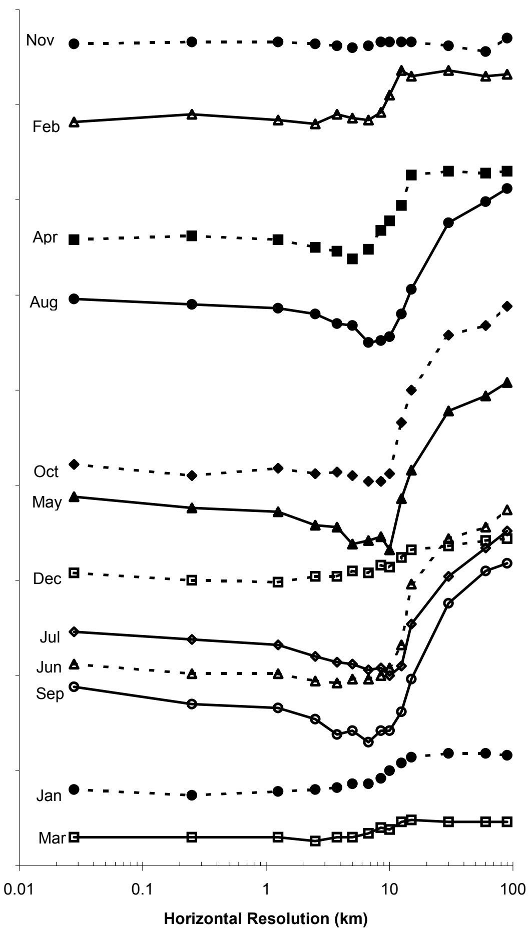


Figure 2. Monthly plots of GCV versus Horizontal Resolution of Topography for the SA region.

In figure 2, the majority of the months show local or global minima in the vicinity of a horizontal resolution of about 5 to 8 kilometres. Moreover, in the cases of January and December, where no well defined local minima is evident, the GCV still displays an increase once the horizontal resolution exceeds approximately 10 kilometres. In figure 3, the GCV curves are slightly noisier yet we again see the well defined local or global minima in the GCV at approximately 5 to 6 kilometres. The months of July and September are notable exceptions to this trend. Note, however, that almost no rain was recorded in these two months; the network means for July and September were 12 mm and 3 mm, respectively, as compared to the mean over all months of 63 mm.

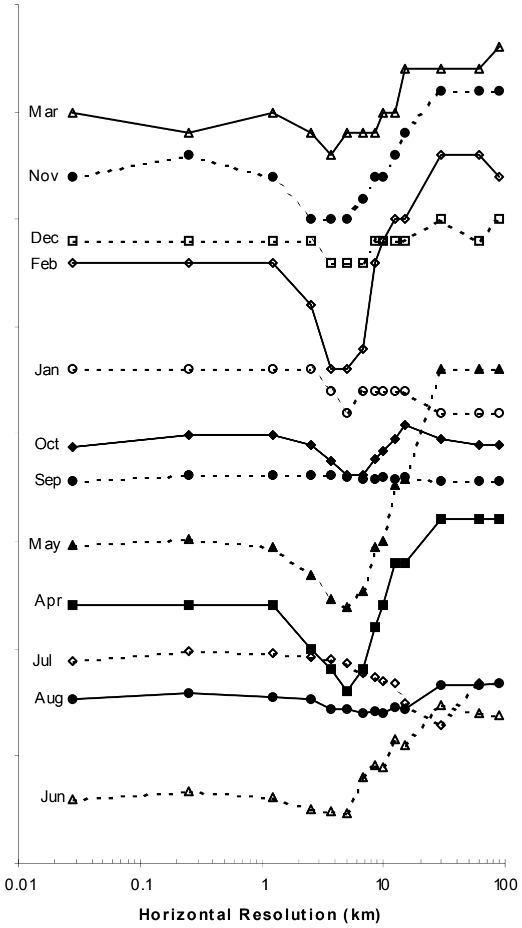


Figure 3. Monthly plots of GCV versus Horizontal Resolution of Topography for the SQ region.

The two regions under study differ in their prevailing climatic conditions. One way this difference manifests itself is in the form of the different seasonal rainfall patterns observed in the two regions. Hence, to better understand what effect this difference might have on the interaction of topography and precipitation, we partition the months into seasonal blocks.

Figure 4 thus takes a slightly more compressed view of how the square root of the GCV varied with the horizontal DEM resolution for the SA region. Shown are the means over the 'winter' months of April through to September and the means over the remaining 'summer' months.

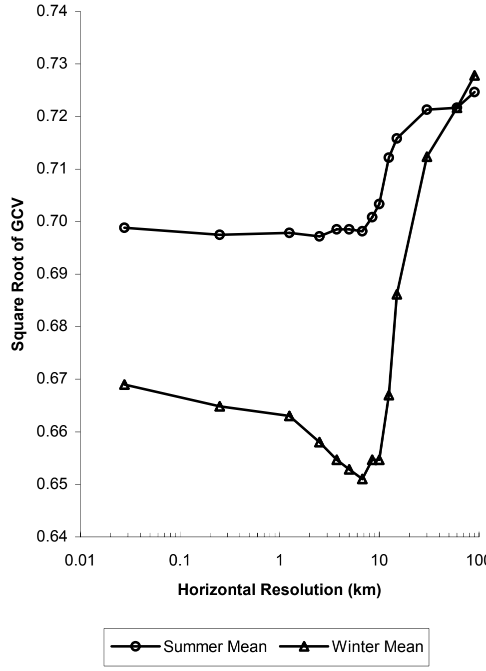


Figure 4. Mean square root of GCV for 'summer' and 'winter' months for the SA region.

The winter mean GCV curve displays a sharp minimum at approximately 7 km resolution while the summer mean GCV curve shows a less defined but similar effect. The reason for the difference between the summer and winter GCV curves is due to the fact that the majority of precipitation in the SA region occurs during the winter months. Thus, the mean rainfall over the data network in July was 57 mm compared to just 5 mm in January. One would not expect to see a significant topographic effect during months when rainfall is slight.

Figure 5 shows the corresponding square root of GCV curves for the SQ region. In contrast to the SA region, both the summer and winter mean GCV curves show definite minima at a horizontal resolution of approximately 5 km. The similar behaviour in summer and winter is due to the relatively high rainfalls that the SQ region receives for the majority of months in the year. The winter rainfall originates predominantly form frontal weather patterns while the summer rainfall is most likely associated with seasonal tropical weather patterns. As mentioned before the only significantly 'dry' months for the SQ region in the year 2000 were July and September. Both of these months elicited almost no response to the resolution of the underlying topography.

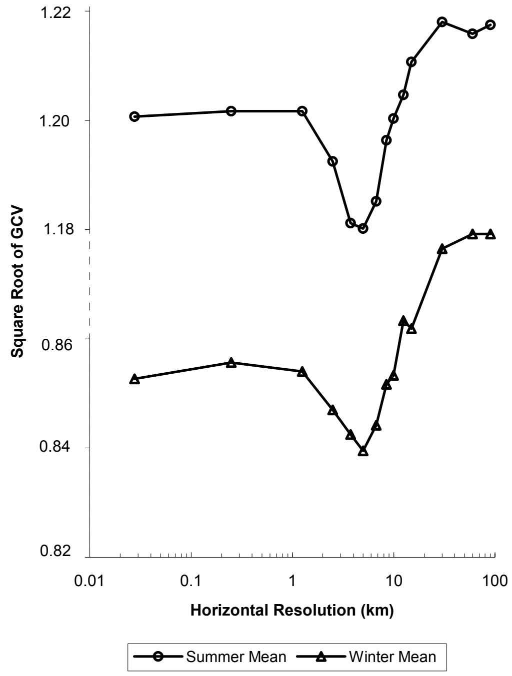


Figure 5. Mean square root of GCV for 'summer' and 'winter' months for the SQ region.

Analogous plots of the validation error for the square root analysis can be seen for the SA region in figure 6 and the SQ region in figure 7. The validation plots are noisier than the GCV plots but in figure 6 the means for both the summer and winter months display broad minima between 5 and 10 kilometres. Again, as in figure 2, the effect is more pronounced in the winter months when the majority of rain is experienced in the SA region.

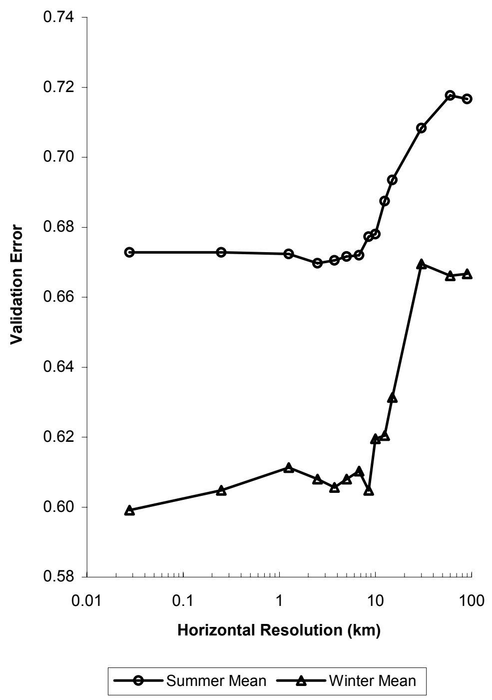


Figure 6. Mean validation error for 'summer' and 'winter' months for the SA region

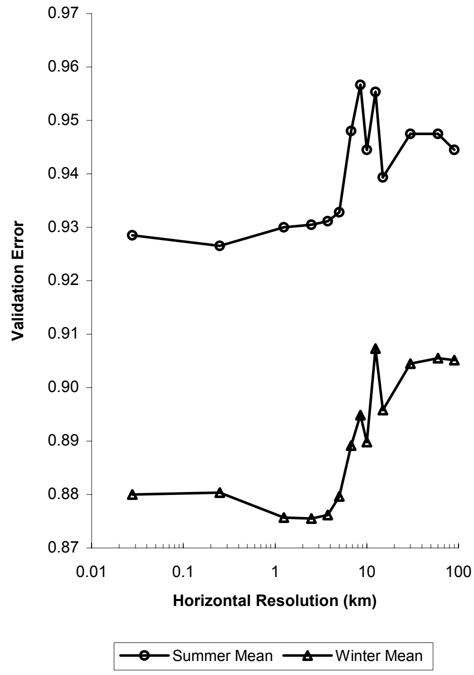


Figure 7. Mean validation error for 'summer' and 'winter' months for the SQ region.

In figure 7 both the summer and winter mean validation error curves behave rather erratically between 8 and 15 kilometres resolution. The winter validation error curve shows a broad minimum at approximately 3 kilometres resolution. This feature, however, is not seen in the summer validation error curve. Given the spatial extent of the SQ region and the sparseness of the data network within it, it would be unwise to place too much emphasis on these results. As well as having regions of significant topography, the SQ region contains relatively flat regions also. If a significant fraction of validation points are found over such flat regions then the ensuing response of the validation errors to horizontal resolution will be corrupted.

5. CONCLUSIONS AND DISCUSSION

This study has shown that the optimal horizontal resolution of topography for the interpolation of monthly rainfall in two subregions of the Australian continent is approximately 5 to 8 kilometres. The fact that this approximate scale suggested itself so predominantly in both regions and in the majority of months considered indicates that 5 to 8 kilometres is a good estimate of the actual scale of the interaction of monthly precipitation with topography in general. Moreover, the result obtained in this paper is in good agreement with the results of Hutchinson [1998] in which the author found an optimal horizontal resolution of approximately 8 kilometres. The precipitation data used in Hutchinson [1998], however, consisted of daily, rather than monthly totals, collected over a portion of the Swiss Alps. The fact that the same resolution was found in both studies suggests that the result is, to an extent, independent of the temporal nature of the precipitation data and the geographical or climatic region in which the precipitation data was collected. This lends further weight to the claim that the atmospheric processes involved in the creation of precipitation respond to the underlying topography at a resolution of approximately 5 to 8 kilometres in general.

This result thus has bearing on the development of atmospheric models. The inclusion in models of a suitable scale of interaction between the atmosphere and topography has the potential to improve model performance and to lead to a better understanding of atmospheric processes.

The result has additional importance in that it suggests that there is a scale of horizontal resolution of topography, below which no real improvement in interpolation error is achieved. Fine scale analyses of environmental processes often require increasingly finer resolution grids of environmental variables and this means that they demand more computer space for their storage. Hence, until improved topographically dependent models are found, it is useful to know that precipitation, despite its spatial complexity, requires only a 5 to 8 kilometre resolution grid of elevation to optimise its topographic dependent interpolation from discrete sources.

The noisy and inconsistent results found for the validation errors in the SQ region are believed to be due to deficiencies in the data network and the relatively large prediction error associated with rainfall validation data sets of modest size. The data network deficiencies are due to the sparseness of the SQ data network and the mixed topographic regimes found within the SQ region.

6. ACKNOWLEDGEMENT

The authors gratefully acknowledge Mr Damian Jellett for preparation of the data sets used in this study and for permission to use his NEARMET routine in their analysis.

7. REFERENCES

Daley, R., *Atmospheric Data Analysis*. Cambridge University Press, New York, 1991.

Houghton, J.T, Meira Filho, L.G., Callander, B.A., Kattenberg, A. and Maskell, K. (Eds), *Climate Change 1995: The Science of Climate Change*. Cambridge University Press, Cambridge, 1996.

Hulme, M., Conway, D., Jones, P.D., Jinang,T., Barrow, E.M., and Turney, C., Construction of a 1961-1990 European climatology for climate change modelling and impact applications. *International Journal of Climatology* 15,1333-1363, 1995.

Hutchinson, M.F. and Bischof, R.J., A new method for estimating the spatial distribution of mean seasonal and annual rainfall applied to the Hunter Valley, New South Wales. Aust. Meteorol. Mag. 31, 179-184, 1983.

Hutchinson, M.F., Stochastic space-time weather models from ground-based data. *Agricultural and Forest Meteorology* 73, 237-264, 1995a.

Hutchinson, M.F., Interpolation of mean precipitation using thin-plate smoothing splines. *International Journal of Geographic Information Systems* 9, 385-403, 1995b.

Hutchinson, M.F., Interpolation of rainfall data with thin-plate smoothing splines II: Analysis of topographic dependence. Journal of Geographic Information and Decision Analysis 2(2), 168-185, 1998.

Hutchinson, M.F., Stein, J.A., and Stein, J.L., *Upgrade of the 9 second Australian digital elevation model*. Centre for Resource and Environmental Studies, Australian National University, 2001. http://cres.anu.edu.au/dem

Hutchinson, M.F., *ANUSPLIN Version 4.3.* Centre for Resource and Environmental Studies, Australian National University, 2003. http://cres.anu.edu.au/outputs/anusplin.html

Hutchinson, M.F., Private Communication, 2002.

Ruddell, A.R., Bidd, W.F., Smith, I.N., Keage, P.L., and Jones, R., *The south-east Australian alpine climate study*. A report by the Meteorology Department, University of Melbourne for the Alpine Resorts Commission, 1990.

Wahba, G., *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM Philadelphia, 1990.