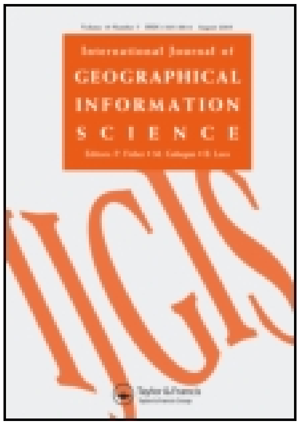
Interpolating Mean Rainfall Using Thin Plate Smoothing Splines

Article *in* Geographical Information Systems · July 1995

DOI: 10.1080/02693799508902045 · Source: DBLP

1 author:

M.F. Hutchinson
Australian National University



International journal of geographical information systems
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tgis19

Interpolating mean rainfall using thin plate smoothing splines

M. F. HUTCHINSON [a]

[a] Centre for Resource and Environmental Studies, Australian National University, GPO Box 4, Canberra, ACT, 2601, Australia E-mail: hutch@cres.anu.edu.au

Available online: 05 Feb 2007

INT. J. GEOGRAPHICAL INFORMATION SYSTEMS, 1995, VOL. 9, NO. 4, 385–403

Research Article

Interpolating mean rainfall using thin plate smoothing splines

M. F. HUTCHINSON
Centre for Resource and Environmental Studies, Australian National University, GPO Box 4, Canberra ACT 2601, Australia
email: hutch@cres.anu.edu.au

Abstract. Thin plate smoothing splines provide accurate, operationally straightforward and computationally efficient solutions to the problem of the spatial interpolation of annual mean rainfall for a standard period from point data which contains many short period rainfall means. The analyses depend on developing a statistical model of the spatial variation of the observed rainfall means, considered as noisy estimates of standard period means. The error structure of this model has two components which allow separately for strong spatially correlated departures of observed short term means from standard period means and for uncorrelated deficiencies in the representation of standard period mean rainfall by a smooth function of position and elevation. Thin plate splines, with the degree of smoothing determining by minimising generalised cross validation, can estimate this smooth function in two ways. First, the spatially correlated error structure of the data can be accommodated directly by estimating the corresponding non-diagonal error covariance matrix. Secondly, spatial correlation in the data error structure can be removed by standardising the observed short term means to standard period mean estimates using linear regression. When applied to data both methods give similar interpolation accuracy, and error estimates of the fitted surfaces are in good agreement with residuals from withheld data. Simplified versions of the data error model, which require only minimal summary data at each location, are also presented. The interpolation accuracy obtained with these models is only slightly inferior to that obtained with more complete statistical models. It is shown that the incorporation of a continuous, spatially varying, dependence on appropriately scaled elevation makes a dominant contribution to surface accuracy. Incorporating dependence on aspect, as determined from a digital elevation model, makes only a marginal further improvement.

1. Introduction

Climate means for months, seasons and years resolve much of the spatial variability of dependent biological activity (Nix, 1986, Booth *et al.* 1987, Mackey *et al.* 1989, Hutchinson *et al.* 1992). Effective determination of the spatial distribution of mean rainfall and other mean climate variables is also a necessary first step towards the development of stochastic models of the weather for more refined assessments of the impacts of climate (Hutchinson 1987) and to provide a basis for spatially interpolating and simulating actual weather values at monthly, weekly and even daily time steps (Hungerford *et al.* 1989, Hutchinson 1995). Climate means also play a crucial role in validating general circulation models (Gates *et al.* 1992). The Weather Generator Project, Focus 4 of Biospheric Aspect of the Hydrological Cycle of the International Geosphere-Biosphere Programme (IGBP-BAHC 1993), is heavily reliant on the development of accurate spatial interpolation methods to disaggregate observed and simulated broad scale climate to the mesoscale. The networks supporting such analyses have been assessed to be critically inadequate on a global basis. A central goal of this paper is to improve the interpolation accuracy that can be obtained from existing data networks. Since the climate stations with records for only small numbers of years often make up the majority in a data network, an important step in this process is to maximise the spatial density of data coverage, and hence interpolation accuracy, by making proper use of rainfall records at all available stations, no matter how short.

Of the standard climate variables, rainfall displays the greatest variability in time and space and so makes the greatest demands on both temporal and spatial analysis techniques. For this reason, rainfall is also the most commonly measured climate variable. Methods for the interpolation of mean rainfall from ground-based point data have ranged from simple trend surface analyses (Edwards 1972, Hughes 1982) and local bi-variate interpolation techniques based on thiessen polygons (Thiessen 1911), inverse distance weighting (Shepard 1968) and delauney triangulations (Akima 1978), through to more sophisticated statistical methods. The last include geostatistical methods (Chua and Bras 1982, Hevesi *et al.* 1992, Phillips *et al.* 1992) and thin plate splines (Hutchinson and Bischof 1983). These methods can produce spatially distributed estimates of interpolation error and the studies cited have all demonstrated the value of incorporating dependence of spatial patterns of rainfall on elevation.

When interpolating rainfall means for a standard period from short records, it is particularly important to take account of the large year to year variability of rainfall. An appropriate statistical model of the spatial distribution of observed rainfall means is therefore presented. The error structure of this model allows for departures of observed rainfall means from standard period means due to missing records, and for deficiencies in the representation of mean rainfall as a smooth function of position and elevation. The first component of the error is due to the relatively large year to year variation of rainfall. The broad scale processes which give rise to this variation imply strong correlations in observed means with overlapping periods of record. This component can therefore exhibit strong dependencies between different data locations, depending on the number of years of common record. The second component of the error is due essentially to local effects below the spatial resolution of the data network. These can be assumed to be independent between different data locations. By simultaneously employing both temporal and spatial dependencies, the method based on this error model is a significant advance over methods, as described by Bennet *et al.* (1984), which estimate missing data by incorporating these dependencies separately.

The non-diagonal covariance structure of the data error model can be accommodated directly, by fitting a spline which uses a corresponding non-diagonal error covariance matrix. Alternatively, the observed short period rainfall means can be standardised to more accurate estimates of the standard period means by adopting the common practice of performing linear regressions with nearby stations. The standardized means can then be interpolated by fitting a spline, with a diagonal error covariance matrix, which allows for spatially varying error variances across the data network. Both methods permit the use of rainfall means recorded over essentially all periods, no matter how short.

The main drawback of these methods is the amount of data required. Annual mean rainfall data for a region in south eastern Australia are used to show that the error structure can be replaced by simpler uncorrelated error structures which give rise to interpolated surfaces of only slightly inferior accuracy. For the simplest of these models, only the rainfall means and their lengths of record are required. Such simple summary data may be the only rainfall data available for some regions.

Thin plate smoothing splines also permit straightforward incorporation of varying degrees of topographic dependence, particularly when they are extended to include linear parametric sub-models. Such functions are called *partial* splines. The paper concludes by examining several spline and partial spline models which incorporate varying degrees of dependence on topography, including both aspect and elevation. The scaling of elevation used in the trivariate smoothing spline function of longitude, latitude and elevation is shown to be critical to its performance.

2. Partial thin plate splines

Partial thin plate splines have been described in detail by Wahba (1990). They include thin plate splines (Wahba and Wendelberger 1980) as a special case. A summary of the basic methodology, with climate interpolation principally in mind, can be found in Hutchinson (1991). Further developments, intended for various geographic applications have been demonstrated by Gu and Wahba (1993) and Mitasova and Mitas (1993). Recent comparisons with geostatistical (kriging) methods, with which splines share close formal connections, have been presented by Hutchinson (1993 a), Hutchinson and Gessler (1994) and Laslett (1994). The main advantage of thin plate splines over competing geostatistical techniques is that splines do not require prior estimation of spatial auto-covariance structure. This structure can be difficult to estimate and validate. Standard variogram fitting techniques may not be used to estimate this structure when, as in virtually all of the analyses below, the variances of the data errors vary across the data network. Moreover, the amount of data smoothing imposed by thin plate splines can be simply optimized by minimizing the generalized cross validation (GCV). This ease of operation does not sacrifice accuracy and the data smoothing paradigm embodied in splines permits efficient detection of data errors (Hutchinson 1993 a, Hutchinson and Gessler 1994). Efficient computational techniques for calculating and optimizing thin plate splines have been developed and are described below.

The partial spline observational model for n data values z_i at positions x_i is given by setting

$$z_i = f(x_i) + \sum_{j=1}^{p} \beta_j \psi_j(x_i) + \varepsilon_i \quad (i = 1, \ldots, n; j = 1, \ldots, p) \tag{1}$$

where f is an unknown smooth function to be estimated, the ψ_j are a set of p known functions and the β_j are a set of unknown parameters which have also to be estimated. The x_i commonly represent coordinates in two or three dimensional euclidean space. The ε_i are zero mean random errors with covariance structure given by

$$E(\varepsilon\varepsilon^T) = V\sigma^2 \tag{2}$$

where $\varepsilon = (\varepsilon, \ldots, \varepsilon_n)^T$, V is positive definite $n \times n$ matrix and σ^2 may be known or unknown. The errors ε_i are uncorrelated if V is diagonal and correlated otherwise. The function f and the parameters β_j are estimated by minimizing

$$(z - g)^T V^{-1} (z - g) + \rho J_m(f) \tag{3}$$

where $z = (z_1, \ldots, z_n)^T$ and $g = (g_1, \ldots, g_n)^T$ with

$$g_i = g(x_i) = f(x_i) + \sum_{j=1}^{p} \beta_j \psi_j(x_i). \tag{4}$$

$J_m(f)$ is a measure of the roughness of the spline function f defined in terms of mth order derivatives of f and ρ is a positive number called the smoothing parameter. The solution to the minimization problem can be solved explicitly, with the estimate of f having an expansion in terms of a scalar function of distance from each data position. The form of the scalar function depends on the dimension of the x_i and the order of derivative m defining the roughness penalty (Wahba 1990). The solution reduces to an ordinary thin plate smoothing spline when there is no parametric sub-model (i.e., when $p = 0$).

The smoothing parameter ρ determines a trade-off between data infidelity and surface roughness. It is usually calculated by minimizing the GCV. This is a measure of the predictive error of the fitted surface which is calculated by removing each data point in turn and summing, with appropriate weighting, the square of the discrepancy of each omitted data point from a surface fitted to all the other data points. It is possible to calculate the GCV implicitly, and hence efficiently. Intuitively, the GCV is a good measure of the predictive power of the fitted surface, as has been verified both theoretically and in applications to real and simulated data (Wahba 1990, Hutchinson and Gessler 1994). The method affords an estimate of σ^2 by analogy with linear regression. Goodness of fit may be assessed by comparing this estimate with *a priori* estimates of σ^2 when they are available.

Computational techniques for determining the smoothing parameter ρ and solving for f and β_j have been optimized for speed and storage space (Hutchinson 1984, Bates *et al.* 1987). The number of data points is normally limited by the requirement to perform a tri-diagonalization of a full matrix with dimension slightly less than n. However, when there are more than a few hundred data points, the thin plate spline function f can be defined in terms of a restricted set of data positions or knots. The computations can then be arranged so that only matrices with order equal to the number of knots have to be stored and decomposed (Hutchinson 1984). This permits data sets with up to a few thousand points to be easily processed on quite modest computers, provided that the fitted spline function can be well described by no more than a few hundred knots. Implementations of these techniques in FORTRAN have been described by Hutchinson (1984) and Bates *et al.* (1987). A software package called ANUSPLIN is available from the author.

3. A serially incomplete annual mean rainfall data set

Annual rainfall data for various subsets of the 34 years between 1955 and 1988 were available at 195 locations for the region of south-eastern Australia shown in figure 1. The spatial density of the data was heavily biased towards the central region of the Australian Capital Territory. The data included those used in a preliminary study by Hutchinson and Johnson (1991). Topographic relief across the entire region plotted in figure 1 was over 1600 m. As shown in figure 2, relatively few stations had complete or even near complete records.

To assess the accuracy of the various interpolation strategies, 18 stations, as also shown in figure 1, were withheld from the surface fitting analyses. These stations were selected, using the SELNOT program in the ANUSPLIN package, to approximately equi-sample the three dimensional space spanned by the data, in which longitude, latitude and elevation were each scaled to have unit variance. The withheld points were selected by initially selecting all data points and then successively removing, from the closest pair remaining in the network, the data point with the shorter period of record, until just 18 points remained. This selection procedure imposed an exacting test on the proposed interpolation methods, since the withheld data points, most of which had more than 30 years of record, would normally be seen as the most desirable to be used to fit

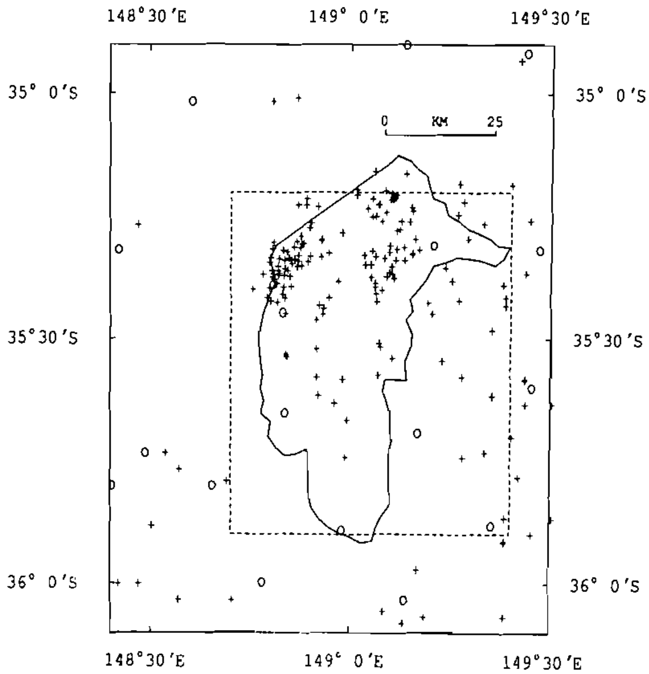


Figure 1. Geographic locations of 195 data points: 177 data points (+) used for surface fitting and 18 data points (O) withheld for validation. The inner solid line delimits the Australian Capital Territory. The dashed line delimits the area plotted in figures 4, 5.

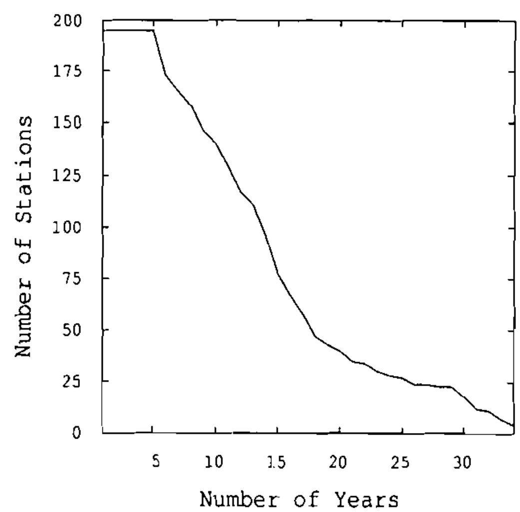


Figure 2. Number of stations for each length of record. Each station had at least 5 years of record.

the surfaces. Some of the withheld stations were quite remote from the other data points, especially in terms of elevation.

3.1. *Standardization to the 34 year period*

It is well known that the annual rainfall mean z_i at location i, calculated from n_i years of record, has variance σ_i^2/n_i, where σ_i^2 is the variance of the annual rainfall at location i. This is based on the assumption that the mean is estimating the 'true' long term mean for an infinitely long period. It is arguable whether such exists, especially when the climate is thought to be changing. A more clearly defined notion is the mean for a specified standard period. It is shown in the Appendix that the data mean at location i, as an estimate of the 34 year mean for the period from 1955 to 1988, has variance

$$(1/n_i - 1/34)\sigma_i^2. \tag{5}$$

This value is zero if the observation record is complete and positive otherwise. The square root of the mean of these values over the complete data set of 195 points was 65·9 mm, about 8 per cent of the network rainfall mean of 870 mm. Underlying this calculation is the reasonable assumption of *weak temporal stationarity*, in the sense that the annual rainfall variance is assumed to be constant only for the duration of the 34 year standard period.

Each observed rainfall mean was converted to a more accurate estimate of the 34 year mean by regressing the observed annual rainfall values on the annual rainfall values at a nearby station with a longer period of record in order to fill in missing values. For each short term rainfall mean, the regression estimate with the smallest estimated error variance was selected. This variance included the variance of the error due to the regression and the variance of the error given by equation (5) when the chosen longer term neighbour itself did not have a complete observation record. The root mean square of the estimated error variances of the standard period estimates, for the selected 177 data points, was 20·6 mm, just 2 per cent of the network mean. Since the 18 withheld data points had longer periods of record, the root mean square of their error variance estimates was 16·5 mm, also about 2 per cent of the network mean.

The standard period regression estimates for the 18 withheld data points and their estimated error variances were used to assess the accuracy of the interpolated thin plate smoothing spline surfaces described below. It should be noted that the withheld data were not used to optimize smoothing parameters, but merely to validate the various error estimates. Estimated surface errors were minimized in each case by minimizing the GCV. In normal applications all 195 data points would be included in the surface fitting process.

4. Two statistical paradigms for interpolating mean rainfall

4.1. *Non-diagonal error covariance model*

A relatively complete statistical model for the observed rainfall means at each location is given by equations (1, 2) with the error structure of the ε_i decomposed into two components by setting

$$V = R + S \tag{6}$$

where R accounts for departures of the observed mean from the 34 year standard period mean and S accounts for deficiencies in the spatial model of mean rainfall given by the smooth function g. It is shown in the Appendix that the elements of R are given by

$$r_{ij} = \left(\frac{n_{ij}}{n_i n_j} - \frac{1}{34}\right) \rho_{ij}\sigma_i\sigma_j \tag{7}$$

where n_i and n_j are the number of years of record at locations i and j, n_{ij} is the number of years in common, ρ_{ij} is the correlation between annual rainfall at each location and σ_i^2 and σ_j^2 are the variances of annual rainfall at each location.

It is reasonable to assume that the deficiencies in the model given by g, due to measurement error and other local effects below the spatial resolution of the data network, are independent from one location to another, and that the magnitude of this error is proportional to the annual rainfall variance at each location. If required, systematic bias in rain gauge readings can be accommodated by first systematically perturbing the observed data. Thus S was assumed to be a diagonal matrix with diagonal elements given by

$$s_{ii} = \sigma_i^2/m \tag{8}$$

where m is an unknown constant factor. The off-diagonal elements of R are zero when either $n_i = 34$ or $n_j = 34$. Thus, the covariance structure is diagonal when all records are complete, and usually non-diagonal when records are incomplete.

The variance σ_i^2 at location i was estimated directly from the annual rainfall data. In order to ensure that the matrix R was positive definite, a parametric form for the inter-station correlation was estimated by fitting to the observed correlations an exponential function of station separation d_{ij} in kilometres given by

$$\rho_{ij} = \exp(-d_{ij}/400). \tag{9}$$

This model of inter-station correlation is displayed in figure 3. A more sophisticated, non-isotropic model which better matched the observed correlations could have been considered, but was not deemed to be necessary due to the very large distance scale of the fitted correlation model and the dominance of the numbers n_i, n_j and n_{ij} in determining the elements of R.

A thin plate spline function of longitude and latitude in degrees and elevation in kilometres was fitted using this non-diagonal error covariance model. The value of the unknown scale factor m was determined by adjusting its value until the value of σ^2 in equation (2) was estimated to be 1·0 by the minimum GCV model, making the relative error structure given by equation (6) absolute. This procedure needs to be examined further, but the results presented in table 1 appear to be reasonable.

The fitted value of m was 80. This indicated that the root mean square of the variances of the departures of the unknown smooth function g from the actual standard period mean annual rainfall at the 18 withheld data points was 26·6 mm, about 3 per cent of the network rainfall mean. The root mean square of the standard errors of the fitted spline estimate of g at the 18 withheld data points was estimated by the method indicated in (Hutchinson 1993 a, Hutchinson and Gessler 1994) to be 45·2 mm. The estimated root mean square validation error given in table 1 was calculated by allowing for these two errors and for the standard errors in the standardized means at the 18 withheld data points given above. The estimated root mean square validation error was thus given by

$$(26{\cdot}6^2 + 45{\cdot}2^2 + 16{\cdot}5^2)^{1/2} = 55{\cdot}0\,\text{mm} \tag{10}$$

which agrees well with the actual root mean square validation residual of 60·5 mm. This is just 7 per cent of the network rainfall mean and in fact less than the root mean square error of the original data means given by 65·9 mm.

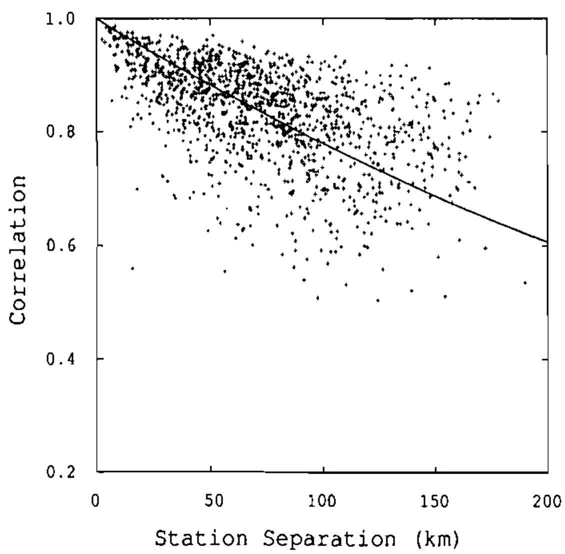


Figure 3. Inter-station correlations versus separation in kilometres.

Table 1. Estimated and actual root mean residuals of 18 data points, withheld from tri-variate spline fits to the remaining 177 data points, using non-diagonal and pre-standardized statistical observational models.

Model	Estimated RMS residual of withheld data (mm)	RMS residual of withheld data (mm)
Non-diagonal error model	55·0	60·5
Pre-standardized data	48·7	62·1

A three-dimensional perspective view of the interpolated annual rainfall overlying a 1/100th degree DEM is shown in figure 4 and a similar view of the estimated standard errors is given in figure 5. Both figures clearly show dependence on elevation. The value of the standard error surface at position x was given by

$$\mathrm{se}\,(x) = (\mathrm{se}\,(g(x))^2 + \sigma^2(x)/80)^{1/2} \tag{11}$$

where $\mathrm{se}\,(g(x))$ was the estimated standard error in estimating $g(x)$ and $\sigma^2(x)$ was the annual rainfall variance at x estimated from the interpolated rainfall mean by the relationship shown in figure 6. This was a linear regression of the logarithms of the observed rainfall variances on the logarithms of the observed rainfall means. The standard errors were thus greatest in high rainfall areas which were remote from the data points. The mean rainfall over the total area shown in figure 1 was determined, using the fitted spline and a 1/100th degree digital elevation model (DEM), to be 907 mm with an estimate standard error of 38 mm, just 4 per cent of the areal mean. The areal mean was larger than the data network mean of 870 mm. Its standard error was estimated using the derived covariance structure of the surface coefficients to estimate the standard error of the areal mean of the estimate of the unknown function g. This assumed, quite reasonably, that the errors in g, which were assumed *a priori*

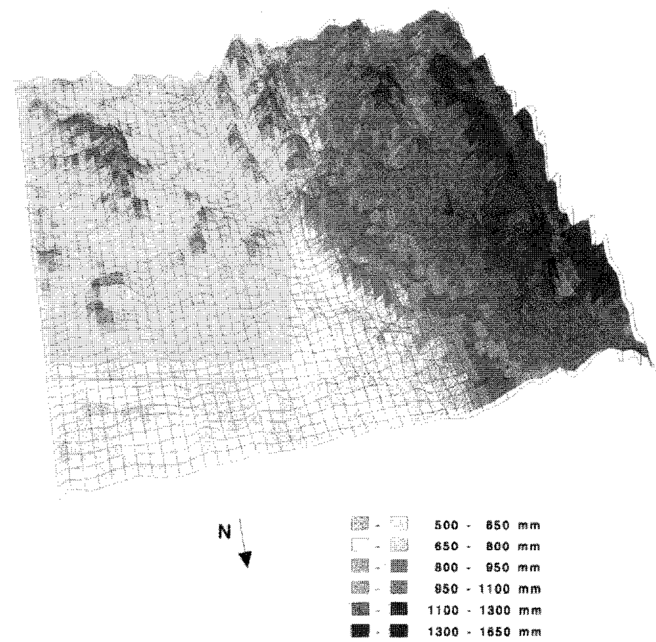


Figure 4. Annual mean rainfall obtained from a tri-variate spline function of longitude, latitude and elevation, with correlated error model, and plotted over a 1/100th degree DEM.

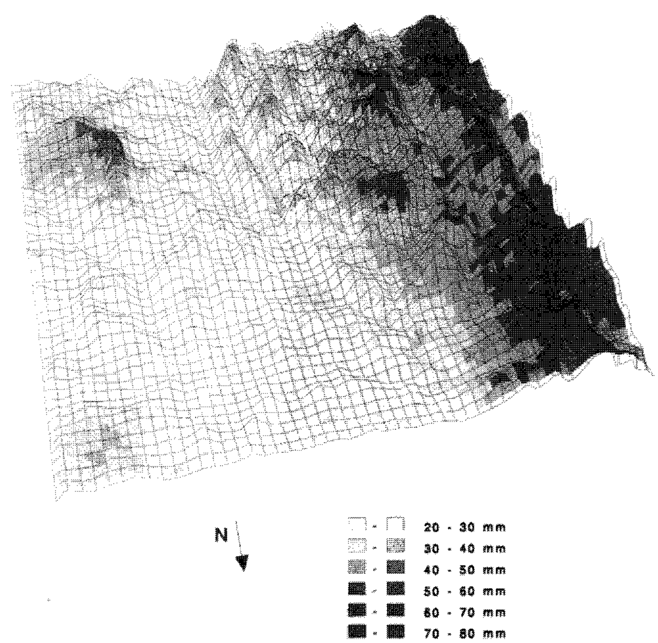


Figure 5. Estimated standard error surface for the annual mean rainfall plotted in figure 4.

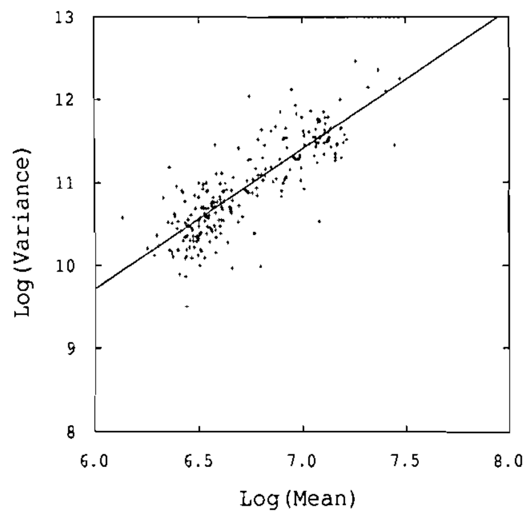


Figure 6. Linear regression of the logarithm of rainfall variance on the logarithm of annual mean rainfall.

to be independent across the data network, approximately cancel when calculating the areal mean.

4.2. *Diagonal pre-standardised model*

An alternative to the above model makes use of the strong inter-station correlations to first standardize the short period means to standard period means using linear regression. As for the non-diagonal procedure, this requires complete knowledge of the actual observations made at each location, but does avoid having to estimate a positive definite error structure for the inter-station correlations. As for the non-diagonal model, it was assumed that the variance of the errors in the unknown function g, which then described the standardized data means, was proportional to the data variance at each location.

The pre-standardized values were thus smoothly interpolated by a minimum GCV thin plate spline function of longitude, latitude and elevation with diagonal relative error structure given by

$$v_{ii} = \sigma_i^2. \tag{12}$$

The minimum GCV estimate of σ^2 in equation (2) effectively estimated the unknown factor m, as in equation (8), and was found to be 50, giving a root mean square standard error across the 18 withheld data points of 33·6 mm. This was larger than the value of 26·6 mm obtained by the non-diagonal procedure, as might be expected, since the diagonal error variance structure had to include the standard errors of the standardized data, given by 16·5 mm. In fact, the root mean square value given by

$$(26{\cdot}6^2 + 16{\cdot}5^2)^{1/2} = 31{\cdot}3\ \text{mm} \tag{13}$$

agreed closely with the root mean square error value of 33·6 mm.

The estimated root mean square errors in the fitted spline estimates of g at the 18 withheld data points was 31·1 mm, rather less than the corresponding value of 45·2 mm obtained for the non-diagonal procedure. The estimated validation root mean square residual in table 1 was then given by

$$(33{\cdot}6^2 + 31{\cdot}1^2 + 16{\cdot}5^2)^{1/2} = 48{\cdot}7\ \text{mm} \tag{14}$$

which was also in agreement with the actual root mean square residual. The actual residuals of the withheld data from the non-diagonal and diagonal models were in very close agreement with each other. The mean rainfall over the area shown in figure 1, determined as above, also agreed closely with the mean for the previous model. It was 914 mm with an estimated standard error of 26 mm (3 per cent of the areal mean).

5. Practical approximate methods

Detailed year by year rainfall records are difficult and sometimes impossible, to obtain. It may also be desirable to avoid detailed calculations of pair-wise regressions using all year by year data, particularly if this does not lead to a significant increase in interpolation error. The statistically precise error models investigated above were replaced by progressively simpler, less precise models. These can be applied when only brief summary data are available at each location. In each case the thin plate smoothing spline was optimized by minimizing the GCV.

5.1. *Weighting by local variance estimates*

In the first simplification, the actual data means were weighted by a diagonal error structure given by

$$v_{ii} = \sigma_i^2/n_i \tag{15}$$

where σ_i^2 was the local variance estimated from the data at location i and n_i was the number of years of record. This was the weighting used by Hutchinson and Bischof (1983). It recognised that stations with higher variances or shorter periods of record should be subjected to more data smoothing than stations with smaller variances or longer periods of record. However, it ignored the strong correlations between means at locations with overlapping periods of record. Some of this spatial correlation can be interpreted as signal by the spline fitting procedure. Nevertheless, the estimated root mean square residuals of the withheld data in table 2 were quite comparable with the actual values. More importantly, the overall performance of the model, as assessed by the validation residual, was only slightly inferior to that of the statistically complete models described above.

Table 2. Estimated and actual root mean square residuals of 18 data points, withheld from tri-variate spline fits to short period means at the remaining 177 data points. Data weighted according to four approximate diagonal error covariance structures.

Model	Estimated RMS residual of withheld data (mm)	RMS residual of withheld data (mm)
Local variance estimate	77·2	72·0
Regression estimate of local variance	79·4	74·2
Length of record	75·8	75·4
Uniform	68·6	76·8

5.2. *Weighting by local variances when variances are unknown*

In some cases variances are not easily obtained. Actual observed variances were then replaced by estimates obtained from observed rainfall means via the simple regression model in figure 6. Though the estimates of variances are not particularly precise, table 2 shows that the model performed virtually as well as the model which used actual variances. This method can also be applied to monthly mean data since reasonably linear relationships are also known to hold between the logarithms of monthly variances and the logarithms of monthly means.

5.3. *Weighting by length of record*

If variances do not vary greatly across the data network, it may be reasonable to weight the observed rainfall means simply by setting

$$v_{ii} = 1/n_i. \qquad (16)$$

Experience in fitting monthly rainfall means across larger areas, for which there can be large spatial variation in rainfall means and variances, suggests that this is not always a viable option. However, it does enjoy the significant practical advantage of giving the same weighting for each month when interpolating monthly mean rainfall. This permits significant saving in computation and subsequent storage of the fitted surface coefficients, since the numerically expensive part of the calculation needs only to be done once in order to fit minimum GCV surfaces to monthly rainfall for each month of the year. For the small regions considered here, the root mean square residual in table 2 was similar to that obtained with the preceding models.

5.4. *Uniform weighting*

Uniform, diagonal weighting is often the default and sometimes only available option for interpolation methods. While this is in fact reasonable for some climate variables such as temperature, for which data means 'settle down' after only a few years (in the absence of longer term climate change), it is not generally recommended for interpolating mean rainfall. However, again because the region considered here is relatively small, the root mean square residual in table 2 was only slightly larger than the values obtained with the preceding models.

6. The relative scaling of elevation

The tri-variate splines described above all used longitude and latitude in degrees and elevation in kilometres. These units were not only convenient but also very appropriate for the analyses considered here. Figure 7 shows a plot of the root mean square validation residual for the tri-variate spline fitted to the pre-standardized data with diagonal weighting. The scaling of elevation was varied through five orders of magnitude from $10^{-2.5}$ to $10^{+2.5}$ kilometres. The validation residual was acceptably small only for values between about 0·5 and 2 kilometres.

7. Incorporating additional topographic effects

It is generally acknowledged that there are significant effects on rainfall related to the aspect. It is also acknowledged that the effects of topography on received rainfall are spatially coarser than the finest descriptions of topography. Both of these effects have been recently examined using local regression methods in the PRISM model (Daly *et al.* 1994). Alternative approaches based on partial thin plate splines were

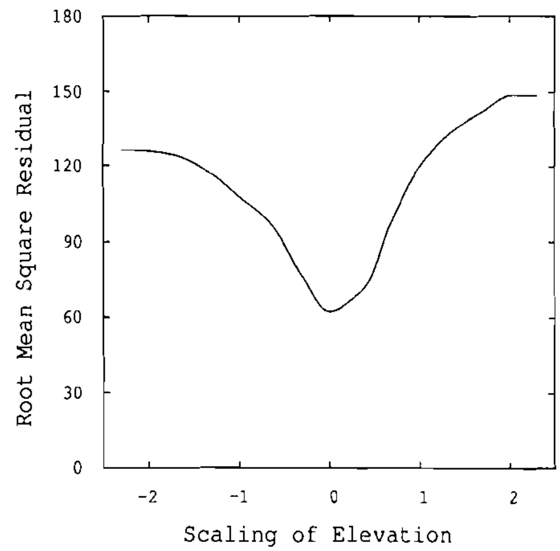


Figure 7. Root mean square residual of data withheld from the minimum GCV tri-variate spline versus scaling of elevation (kilometres) in powers of 10.

examined here. Using the pre-standardized data described above, partial thin plate splines, with increasingly complex dependencies on topography, were fitted as follows:

(*a*) Bi-variate thin plate spline function of longitude and latitude only.
(*b*) Tri-variate partial thin plate spline incorporating a bi-variate thin plate spline function of longitude and latitude and a constant linear dependence on elevation obtained from a 1/40th degree DEM (Hutchinson 1993 b).
(*c*) Tri-variate thin plate spline function of longitude, latitude and elevation obtained from a 1/40th degree DEM.
(*d*) Quint-variate partial thin plate spline incorporating a tri-variate thin plate spline function of longitude, latitude and elevation derived from a 1/40th degree DEM and a linear dependence on the two horizontal components of the unit normal vector on the same 1/40th degree DEM.

The results are presented in table 3. The bi-variate thin plate spline and the tri-variate partial spline were plainly inferior, both in terms of actual and estimated root mean square residuals, to the splines which incorporated a spatially varying dependence on elevation. The mean rainfall over the area covered in figure 1 was estimated by the

Table 3. Estimated and actual root means square residuals of 18 data points, withheld from four spline models fitted to pre-standardized data means at the remaining 177 data points, with varying degrees of topographic dependence on a 1/40th degree DEM.

Model	Estimated RMS residual of withheld data (mm)	RMS residual of withheld data (mm)
Bi-variate spline	86·9	179·5
Tri-variate partial spline	87·5	139·9
Tri-variate spline	41·6	57·2
Quint-variate partial spline	40·2	56·4

bi-variate spline to be 881 mm with a standard error of 52 mm (6 per cent). Unlike the mean value estimated by the tri-variate spline, this mean was similar to the network mean rainfall. The estimated standard error was also larger than the standard errors estimated from the tri-variate splines.

The tri-variate thin plate spline function, which used elevations from the 1/40th degree DEM, performed similarly to the tri-variate spline function derived from actual station elevations (see table 1). The quint-variate partial spline was only marginally superior to the tri-variate thin plate spline. The fitted coefficients of the eastern and northern components of the unit normal vector were 113 and 27 mm/m respectively, with respective estimated standard errors of 55 and 80 mm/m (see Hutchinson 1993 a). This implied that the rainfall means were perturbed positively from a relatively simple function of elevation for areas facing the neighbouring east coast of Australia and perturbed negatively for areas facing away from the coast. By basing the effect on the components of the unit normal, rather than the two components of aspect, the effects were more pronounced in areas with steeper slopes. These components were also continuous functions of position on the DEM.

8. Discussion

Partial thin plate splines have been shown to be flexible and computationally efficient tools for interpolating mean rainfall for a standard period. They were able to incorporate the correlated error structure inherent in rainfall means obtained from incomplete records and provided accurate spatially distributed estimates of the error of the interpolated surfaces, as verified using withheld data. For the two complete data error models, these errors were just 7 per cent of the network mean and the areal mean rainfall was estimated with standard error estimates of 3–4 per cent. The spatially distributed error estimates were continuous and took account of both station density and larger variability in higher rainfall regions. Such estimates are not routinely available for either the PRISM method (Daly *et al.* 1994) or the MTCLIM method (Hungerford *et al.* 1989). However, it should be noted that MTCLIM was designed for interpolating actual rainfall values once underlying annual mean values were obtained from an isohyetal map. This paper is principally concerned with the interpolation of such underlying means.

Since the PRISM and MTCLIM methods are both based on developing detailed local empirical regression relationships between rainfall and topography, they do not yield a continuous interpolated rainfall field. In consequence, the PRISM method employs *post hoc* filtering and other adjustments to produce an acceptable interpolated field. Tri-variate thin plate splines naturally allow for a continuous spatially variable dependence on elevation and are therefore well suited for application across large heterogeneous areas (Hutchinson 1991, 1993 b), answering a call for such methods by Dolph and Marks (1992). Tri-variate splines impose strong continuity conditions on this elevation dependence, making it robustly determined from scattered and moderately sparse data networks. This has rendered insignificant the contributions of adding further dependence on generalized aspect obtained from a DEM. Neither were there significant differences between using actual station elevations and using elevations on a relatively coarse DEM. Nevertheless, the appropriate spatial scale of elevation and aspect effects clearly deserves further investigation. For the data analysed here, the marginal benefits obtained by considering broad scale DEM elevation and aspect were clearly outweighed by the benefits obtained by correctly scaling elevation.

The model which required standardization of observed short period means to standard period estimates had the same data requirements as the model which used the non-diagonal error covariance structure, but was otherwise simpler to use, since it avoided having to estimate and store a full, positive definite, error covariance matrix. Once the data had been standardized they were able to be processed by methods which permitted only diagonal weighting. This can be particularly useful when processing very large data sets with the approximate version of thin plate splines defined in terms of knots.

The simpler non-correlated error structure used by Hutchinson and Bischof (1983) yielded only a modest reduction in the accuracy obtained by the more complete error models. The obvious practical advantages of this structure were that it depended on simply summary data only at each location and that it further reduced computational requirements. It also gave an indication of the robustness of thin plate smoothing splines when error variance structure is not accurately specified.

In keeping with the results of Chua and Bras (1982) and Phillips *et al.* (1992), the incorporation of a constant linear dependence on elevation produced a more accurate result than having no elevation dependence. However, errors were much more significantly reduced by using a tri-variate thin plate spline which incorporated a continuous spatially varying dependence on elevation. Previous attempts to incorporate more complex dependencies on elevation using *cokriging* have not led to significant improvements over methods with constant linear dependence on elevation (Chua and Bras 1982, Phillips *et al.* 1992). The larger estimated and actual validation residuals obtained with surfaces which did not incorporate a spatially varying dependence on elevation clearly indicated the unreliability of such methods, especially in areas remote from the data points.

A important practical advantage of using actual station elevations when fitting thin plate splines is that access to a DEM was not required, neither when fitting the surface nor when calculating fitted values at particular points. A DEM was only required when producing images such as those in figures 4 and 5 and in calculating areal mean rainfall. The fitted rainfall surface can be stored compactly as a set of surface coefficients describing the fitted function of longitude, latitude and elevation. Files of these coefficients can be used to construct a readily accessible climate surface data base. Such climate surfaces have already been constructed for whole continents, including Australia (Hutchinson 1991, 1993b) and Africa (Hutchinson *et al.* 1993), from continent-wide climate data consisting of thousands of points.

The production of these climate surfaces has required only the computational resources afforded by a modest workstation. The activity which has demanded the majority of both user and computer time has been the compilation of large accurately geolocated data sets. This has been significantly facilitated by the efficient error detection procedure afforded by thin plate smoothing splines in which erroneous data, particularly those data with errors in geographic position or elevation, are indicated by large outliers from the fitted spline surface.

9. Summary recommendations

The recommended options when interpolating standard period mean rainfall surfaces using thin plate smoothing splines depend on the temporal and spatial coverage of the data. In all cases a tri-variate spline which incorporates a spatially varying dependence on appropriately scaled elevation should be applied to all available short period means in order to maximize the spatial density of the data. When using position coordinates in degrees of longitude and latitude, elevation may be scaled in kilometres.

This scaling can be slightly refined by choosing the elevation scaling to minimize the GCV of the fitted surface, but in the author's experience, further error reduction achieved in this way is marginal. It remains to be demonstrated whether incorporating any topographic parameter other than actual station elevation, can contribute significantly to interpolation accuracy, especially when data sets are reasonably dense. Using actual station elevations has the added practical advantage of not requiring a DEM when fitting the thin plate spline.

When year by year data making up each observed mean are available, the simplest procedure appears to be to fit a spline, with a suitable diagonal error covariance structure, to estimates of standard period mean rainfall obtained from observed short period means via linear regression with nearby longer period stations. This also permits the estimation of accurate spatially distributed standard errors. The accuracy of this procedure was essentially the same as that obtained with the procedure which employed the full correlated error structure of the observed rainfall means.

When only summary mean and variance rainfall data are available, or when the spatial density of the data is high, then one of the approximate diagonal error structures listed in Table 2 may be applied to raw short period rainfall means, with only a modest reduction in interpolation accuracy. Weighting by length of record is a practical alternative when rainfall variances do not vary greatly across the region being analysed. It has the added advantage of significantly reducing computational requirements when fitting 12 monthly mean rainfall surfaces. Fitting a spline with uniform weighting to raw short period rainfall means is not generally recommended.

When data sets contain between several hundred and a few thousand points, the approximate procedure using selected knots may be used with little loss in accuracy. For still larger data sets, the region should be divided into smaller pieces. Thin plate splines may be fitted to these pieces, with generous inclusion of data adjacent to each piece in order to remove edge effects.

Acknowledgment

The author gratefully acknowledges the assistance of Jennifer Kesteven in the preparation of the colour figures.

Appendix

The variance and covariance structure of observed short period means, considered as estimates of the standard 34 year mean, are derived.

Variance

Let N denote the set of $n = 34$ years from 1955 to 1988 and let $\{y_i, i \in N\}$ denote the complete set of observations for all n years at one location, with mean $\bar{y}_n$ and variance σ^2. Assume values for each year are independent. Let M denote an arbitrary non-empty subset of N containing m years and let $\bar{y}_m$ denote the mean of the y_i, with i contained in M. The variance of $\bar{y}_m$ as an estimate of $\bar{y}_n$ is given by

$$\operatorname{var}(\bar{y}_m - \bar{y}_n) = \operatorname{var}\left(\frac{\sum_{i \in M} y_i}{m} - \frac{\sum_{i \in N} y_i}{n}\right)$$

$$= \operatorname{var}\left(\frac{\frac{n-m}{m}\sum_{i \in M} y_i - \sum_{i \notin M} y_i}{n}\right)$$

$$= \left(\frac{\left(\frac{n-m}{m}\right)^2 m + n - m}{n^2} \right) \sigma^2$$

$$= \left(\frac{1}{m} - \frac{1}{n} \right) \sigma^2.$$

Covariance

Now let $\{y_i, i \in M\}$ and $\{z_j, j \in P\}$ denote sets of observations at two locations for two arbitrary non-empty subsets M and P of N with m and p elements respectively. Suppose that the y_i and z_j have means $\bar{y}_m$ and $\bar{z}_p$ respectively. Suppose also that the data variances are σ_y^2 and σ_z^2 respectively and that the correlation between the two sets of observations is ρ. Thus, the covariance of two individual observations y_i and z_j is $\rho\sigma_y\sigma_z$ if $i = j$ and 0 otherwise. Let the number of years of common record be

$$n_{yz} = |M \cap P|.$$

Arguing along similar lines to the above, the covariance of $\bar{y}_m$ and $\bar{z}_p$, considered as estimates of the standard period means $\bar{y}_n$ and $\bar{z}_n$, is then given by

$$\frac{\operatorname{cov}(\bar{y}_m - \bar{y}_n, \bar{z}_p - \bar{z}_n)}{\rho\sigma_y\sigma_z} = \frac{\operatorname{cov}\left(\frac{n-m}{m} \sum_{i \in M} y_i - \sum_{i \notin M} y_i, \frac{n-p}{p} \sum_{j \in P} z_j - \sum_{j \notin P} z_j\right)}{\rho\sigma_y\sigma_z n^2}$$

$$= \frac{\left(\frac{n-m}{m}\right)\left(\frac{n-p}{p}\right)|M \cap P| - \left(\frac{n-m}{m}\right)|M \backslash P| - \left(\frac{n-p}{p}\right)|P \backslash M| + |N \backslash M \backslash P|}{n^2}$$

$$= \frac{(n-m)(n-p)n_{yz} - (n-m)(m-n_{yz})p - (n-p)(p-n_{yz})m + (n-m-p+n_{yz})pm}{mpn^2}$$

$$= \frac{n_{yz}}{mp} - \frac{1}{n}.$$

References

Akima, H., 1978, A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points. *ACM Transactions of Mathematical Software*, **4,** 148–159.

Bates, D., Lindstrom, M., Wahba, G., and Yandell, B., 1987, GCVPACK—routines for generalised cross validation. *Communications in Statistics B—Simulation and Computation*, **16,** 263–297.

Bennett, R. J., Haining, R. P., and Griffith, D. A., 1984, The problem of missing data on spatial surfaces. *Annals of the Association of American Geographers*, **74,** 138–156.

Booth, T. H., Nix, H. A., Hutchinson, M. F., and Jovanovic, T., 1987, Niche analysis and tree species introduction. *Forest Ecology and Management*, **23,** 47–59.

Chua, S.-H., and Bras, R. L., 1982, Optimal estimators of mean-areal precipitation in regions of orographic influence. *Journal of Hydrology*, **57,** 23–48.

Daly, C., Neilson, R. P., and Phillips, D. L., 1994, A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, **33,** 140–158.

Dolph, J., and Marks, D., 1992, Characterising the distribution of observed precipitation and runoff over the continental United States. *Climate Change*, **22,** 99–199.

EDWARDS, K. A., 1972, Estimating areal rainfall by fitting surfaces to irregularly spaced data. *Proceedings of the International Symposium on the Distribution of Precipitation in Mountainous Areas 2* (Geneva: World Meteorological Organisation), No. 326, pp. 565–587.

GATES, W. L., MITCHELL, J. F. B., BOER, G. J., CUBASCH, U., and MELESHKO, V. P., 1992, Climate modelling, climate prediction and model validation. In *Climate Change 1992, The Supplementary Report to the IPCC Scientific Assessment*, edited by J. T. Houghton, B. A. Callader and S. K. Varney (Cambridge: Cambridge University Press) 97–134.

GU, C., and WAHBA, G., 1993, Semiparametric analysis of variance with tensor product thin plate splines. *Journal of the Royal Statistical Society, Series* B, **55**, 353–368.

HEVESI, J. A., ISTOK, J. D., and FLINT, A. L., 1992, Precipitation estimation in mountainous terrain using multivariate geostatistics. Part I: Structural analysis. *Journal of Applied Meteorology*, **31**, 661–676.

HUGHES, D. A., 1982, The relationship between mean annual rainfall and physiographic variables applied to the coastal region of southern Africa. *South African Geographic Journal*, **64**, 41–50.

HUNGERFORD, R. D., NEMANI, R. R., RUNNING, S. W., and COUGHLAN, J. C., 1989, MTCLIM: a mountain microclimate simulation model. Research Paper INT-414. Ogden, UT: US Department of Agriculture, Forest Service, Intermountain Research Station. 52 pp.

HUTCHINSON, M. F., 1984, *Some Surface Fitting and Contouring Programs for Noisy Data*. CSIRO Division Mathematics and Statistics Consulting Report ACT 84/6, 24 pp.

HUTCHINSON, M. F., 1987, Methods of generating weather variables. In *Agricultural Environments: Characterisation, Classification and Mapping*, edited by A. H. Bunting (Wallingford: CAB International), 149–157.

HUTCHINSON, M. F., 1991, Continent-wide data assimilation using thin plate smoothing splines. In *Data Assimilation Systems*, edited by J. D. Jasper (BMRC Research Report No. 27, Melbourne: Bureau of Meteorology), 104–113.

HUTCHINSON, M. F., 1993 a, On thin plate splines and kriging. In *Computing and Science in Statistics, Vol. 25*, edited by M. E. Tarter and M. D. Lock (University of California, Berkeley: Interface Foundation of North America).

HUTCHINSON, M. F., 1993 b, Development of a continent-wide DEM with applications to terrain and climate analysis. In *Environmental Modeling with GIS*, edited by M. F. Goodchild, B. O. Parks and L. T. Steyaert (Oxford: Oxford University Press), 392–399.

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

HUTCHINSON, M. F., and BISCHOF, R. J., 1983, A new method for estimating mean seasonal and annual rainfall for the Hunter Valley, New South Wales. *Australian Meteorological Magazine*, **31**, 179–184.

HUTCHINSON, M. F., and JOHNSON, C. M., 1991, Accounting for strong spatial correlations in the interpolation of mean rainfall for a standard period. In *Proceedings of the Conference on Agricultural Meteorology* (Melbourne: Bureau of Meteorology), pp. 315–318.

HUTCHINSON, M. F., and GESSLER, P. E., 1994, Splines—more than just a smooth interpolator. *Geoderma*, **62**, 45–67.

HUTCHINSON, M. F., NIX, H. A., and MCMAHON, J. P., 1992, Climate constraints on cropping systems. In *Ecosystems of the World: Field Crop Ecosystems*, edited by C. J. Pearson (Amsterdam: Elsevier), 37–58.

HUTCHINSON, M. F., NIX, H. A., MCMAHON, J. P., and ORD, K. D., 1993, *Geographic Information Systems for Africa*. Centre for Resource and Environmental Studies Annual Report, Australian National University, Canberra, pp. 2–3.

IGBP-BAHC, 1993, *BAHC News*. No. 1, November 1993. Newsletter of biospheric aspects of the hydrological cycle (BAHC), Core Project of the International Geosphere-Biosphere Programme (IGBP), Institut für Meteorologie, Berlin, 8 pp.

LASLETT, G. M., 1994, Kriging and splines: an empirical comparison of their predictive performance in some applications. *Journal of the American Statistical Association*, **89**, 391–409.

MACKEY, B. G., NIX, H. A., STEIN, J. A., and CORK, S. E., 1989, Assessing the representativeness of the wet tropics of Queensland World heritage property. *Biological Conservation*, **50**, 279–303.

Mitasova, H., and Mitas, L., 1993, Interpolation by regularized spline with tension: I. theory and implementation. *Mathematical Geology*, **25**, 641–655.

Nix, H. A., 1986, A biogeographic analysis of Australian elapid snakes. In *Atlas of Australian Snakes*, edited by R. Longmore (Canberra: Bureau of Flora and Fauna), 4–15.

Phillips, D. L., Dolph, J., and Marks, D., 1992, A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain. *Agricultural and Forest Meteorology*, **58**, 119–141.

Shepard, D. L., 1968, A two dimensional interpolation function for irregularly spaced data. *Proceedings of the 23rd National Conference, Association of Computing Machinery*, (Washington: ACM), pp. 517–524.

Thiessen, A. H., 1911, Precipitation averages for large areas. *Monthly Weather Review*, **39**, 1082–1084.

Wahba, G., and Wendelberger, J., 1980, Some new mathematical methods for variational objective analysis using splines and cross validation. *Monthly Weather Review*, **108**, 122–1143.

Wahba, G., 1990, Spline models for observational data. In *CBMS-NSF Regional Conference Series in Applied Mathematics*, (Philadelphia: Society for Industrial and Applied Mathematics), 169 pp.