ELSEVIER

Agricultural and Forest Meteorology 84 (1997) 255–271

AGRICULTURAL AND FOREST METEOROLOGY

Application of geostatistics to evaluate partial weather station networks

Muhammad Ashraf [a], Jim C. Loftis [a,*], K.G. Hubbard [b]

[a] *Department of Chemical and Bioresource Engineering, Colorado State University, Fort Collins, CO 80523, USA*

[b] *Department of Agricultural Meteorology, University of Nebraska, Lincoln, NE 68583-0728, USA*

Received 6 February 1995; accepted 29 April 1996

Abstract

Climatic data are an essential input for the determination of crop water requirements. The density and location of weather stations are the important design variables for obtaining the required degree of accuracy of weather data. The planning of weather station networks should include economic considerations, and a mixture of full and partial weather stations could be a cost-effective alternative. A 'full' weather station is defined here as one in which all the weather variables used in the modified Penman equation are measured, and a 'partial' weather station is one in which some, but not all, weather variables are measured. The accuracy of reference evapotranspiration (Et_r) estimates for sites located some distance from surrounding stations is dependent on measurement error, error of the estimation equation, and interpolation error. The interpolation error is affected by the spatial correlation structure of weather variables and method of interpolation. A case-study data set of 2 years of daily climatic data (1989–1990) from 17 stations in the states of Nebraska, Kansas, and Colorado was used to compare alternative network designs and interpolation methods. Root mean squared interpolation error (RMSIE) values were the criteria for evaluating Et_r estimates and network performance. The kriging method gave the lowest RMSIE, followed by the inverse distance square method and the inverse distance method. Co-kriging improved the estimates still further. For a given level of performance, a mixture of full and partial weather stations would be more economical than full stations only.

Keywords: Geostatistics; Weather station networks; Crop water network; Climatic data

1. Introduction

Irrigation scheduling is defined as the procedure for predicting the timing and amount of the next irrigation. Scheduling plays an important role in planning for the future of agriculture, especially in developing countries, where resources are often severely limited. There are a number of ways to perform irrigation scheduling. One of the most common approaches uses estimated crop water requirements based on computed reference evapotranspiration (Et_r).

Several definitions of Et_r have been formulated. Jensen (1973) defined Et_r as the rate at which water, if available, would be removed from the soil and plant surface, expressed as the rate of latent heat transfer per square centimeter of surface or alterna-

* Corresponding author.

0168-1923/97/$17.00 © 1997 Elsevier Science B.V. All rights reserved.
PII S0168-1923(96)02358-1

tively as a depth of water per unit time. For comparative purposes it refers to a well-watered crop such as alfalfa (lucern) with 30–50 cm of top growth and about 100 m of fetch under given climatic conditions unless defined otherwise. Duke et al. (1985) simplified the definition of Et_r to 'the water used by a well-watered reference crop, such as alfalfa, which fully covers the soil surface'. The older term, potential Et, refers to water use by any well-watered crop and is, therefore, not defined precisely. Et_r can be computed by using weather data and any of several estimating equations. The availability of weather data of acceptable spatial resolution for large-scale irrigation scheduling is an important factor to consider in planning the development and management of irrigation systems throughout the world.

The modified Penman combination equation is used to compute Et_r, as it is considered to be a satisfactory estimation equation when daily estimates of Et_r are desired (Jensen et al., 1990). The procedure for computing Et_r using this equation is described in Appendix A.

2. Background theory of partial weather stations

The planning of climate station networks has occupied researchers for many years and is becoming even more essential because of resource limitations and the requirement for optimizing the use of climatological data. Perhaps the most famous work in this field was that by Gandin (1970). Network design using only 'full' weather stations has generally been the practice in the recent past. However, it might be more economical to use some mixture of 'full' and 'partial' weather stations, i.e. stations of different classes. Cost considerations might be especially important in developing countries, which face extreme pressures on both their water resources and financial resources.

A 'full' weather station is defined here as one in which all the weather variables used in the modified Penman equation are observed or measured, and a 'partial' weather station is one in which some but not all weather variables are observed or measured. Density and location of different classes of weather stations in a given region have received very little or no attention from researchers. This suggests the importance of the proposed research on (1) design of weather monitoring network density, (2) evaluating the advantages of 'partial' weather stations for accurate measurement of reference evapotranspirtation (Et_r) and (3) location of different classes of weather stations. A number of studies have been conducted in the field of spatial analysis of weather data and network design. Amegee (1985) used kriging to develop regional contour maps for seasonal Et_r. Harcum and Loftis (1987b) worked on a study of weather monitoring network design and evaluation, extending the use of Kalman filtering for selecting the station density and location. That study explicitly considered both spatial and temporal correlation. In the present study only spatial correlation is considered for design purposes, therefore a shorter period of record is needed to characterize the correlation structure of the weather variables. Hubbard (1994) studied spatial variability of daily weather variables in the High Plains region of the central USA.

Gandin (1970) mentioned that the basic meteorological elements traditionally observed at a ground level station can, as a first approximation, be divided into three groups corresponding to the density of the observational network they require. The first group of weather elements, including air pressure, soil temperature at depth and sunshine duration (solar radiation included), requires the least density of networking. The second group, of intermediate spatial density, includes air temperature, humidity, wind speed and cloud cover. The third group, and that of highest spatial density, comprises precipitation, the characteristics of snow cover and also elements covered by the term 'meteorological phenomena' (thunder storms, fog, snow storms). Gandin (1970) also recommended in general, for the flat areas of the USSR, the admissible distances between observation points or stations for each group. The recommended maximum distance between points for least dense networks was 150–200 km, for the intermediate network, 50–60 km and for the most dense network, 30 km.

Some of the weather elements observed in these classes recommended by Gandin (1970) are beyond the scope of this study, and he did not work on the estimation of evapotranspiration. In the present study two classes of weather stations have been defined for network design. They are given below with the weather elements, required for estimating Penman Et_r, observed in each class. This classification is based on the investment and operational cost of the weather stations and the spatial variability of the weather elements. Class B has maximum and minimum temperatures, relative humidity and wind run, and Class A comprises all the elements in Class B and solar radiation. Usually the weather elements observed in Class B require less sophisticated and lower-cost equipment (with the exception of relative humidity) and are of (intuitively) higher spatial variability. (This has been verified using the data from the case study.)

3. Case study

The primary case study data set includes 17 weather stations in the states of Nebraska, Kansas and Colorado. For this study, 2 years of daily climatic data (1989–1990) are selected from these 17 stations. Specific geographical information for the study area is given in Table 1. The H and V points in Table 1 are respectively horizontal and vertical coordinates of a station from a reference point. The reference point is 37:30°N, 104:00°W. The V distances are oriented north and south, and the H distances are oriented east and west. The location of the case study area is shown in Fig. 1.

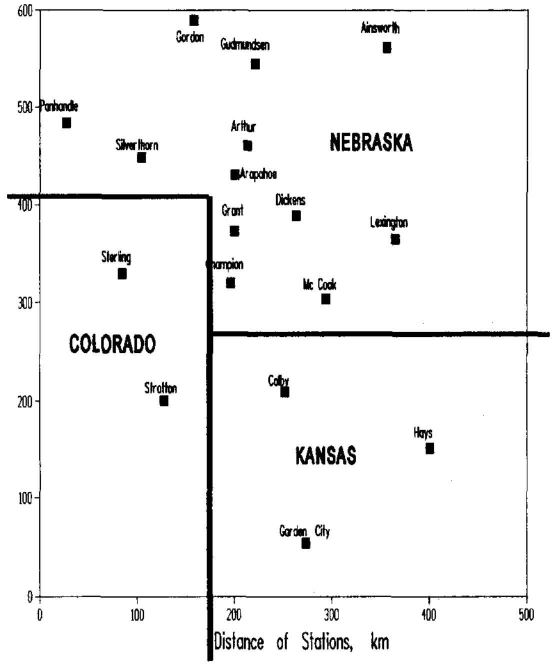


Fig. 1. Study area showing weather stations.

Table 1
Geographical information of case study area

Weather station	Elevation (m)	Latitude (°N)	Longitude (°W)	H (km)	V (km)
McCook, NE	792	40.23	100.58	293.83	303.78
Champion, NE	1029	40.38	101.72	196.36	320.45
Grant, NE	975	40.85	101.67	200.67	372.32
Silverthorn, NE	1302	41.53	102.78	104.63	448.26
Gudmundsen, NE	1049	42.4	101.43	220.73	544.59
Lexington, NE	731	40.78	99.75	365.5	364.91
Dickens, NE	945	41	100.93	263.73	389
Arthur, NE	1097	41.65	101.52	213.57	461.23
Arapahoe, NE	1097	41.37	101.67	200.67	429.74
Panhandle, NE	1244	41.85	103.68	27.23	483.46
Ainsworth, NE	765	42.55	99.87	355.47	561.26
Gordon, NE	1109	42.8	102.17	157.67	589.04
Garden City, KS	866	37.98	100.82	273.77	53.72
Colby, KS	966	39.38	101.07	252.27	209.31
Hays, KS	610	38.87	99.33	401.33	151.89
Stratton, CO	1390	39.3	102.52	127.57	200.05
Sterling, CO	1200	40.47	103.17	84.57	329.72

4. Description of the available data

The data used in this study are from the Automated Weather Data Network (AWDN) operated by the High Plains Climate Center (HPCC) located at the University of Nebraska, Lincoln. The weather elements are observed at the field stations in the case study area (Hubbard et al., 1983). The final data analyzed are the daily, 3 day and weekly averages of the following weather elements for the given station, as required for the modified Penman equation described in Appendix A: maximum and minimum temperature (°C); relative humidity (%); wind run (km day^{-1}); solar radiation (MJ m^{-2} day^{-1}); average computed Et_r (mm day^{-1}).

5. Geostatistical methods used

In the present study mean values were removed from the data to avoid the effect of trend (drift) in developing spatial structures of the weather variables. The following spatial functions were used in this study.

5.1. Semi-variogram

Perhaps the best known spatial function is the semi-variogram, which is a function or graph describing the expected squared difference between pairs of samples with a given relative orientation (Clark, 1982). The general equation for the sample semi-variogram $\gamma(d_{ij})$ of any phenomenon is given as

$$\gamma(d_{ij}) = \frac{1}{2n}\sum_{1}^{n}(X_i - X_j)^2 \tag{1}$$

where the sum is taken over all points separated by distance d_{ij} (actually a distance class interval). The following equation represents the sample semi-variogram after the annual mean (drift or trend) is removed:

$$\gamma(d_{ij}) = \frac{1}{2n}\sum_{1}^{n}\left[\left(X_i - \bar{X}_i\right) - \left(X_j - \bar{X}_j\right)\right]^2 \tag{2}$$

5.2. Spatial correlation

Correlation is a measure of linear dependence between two or more variables, and spatial correlation is the relation between the values of a given variable at a point and the values of the same variable or a different variable (spatial cross-correlation) at a stated distance from the first point. In other words, the spatial correlation is the relationship between one or more variables with a given relative orientation. A spatial correlation function gives values of spatial correlation coefficients (the familiar Pearson's r) as a function of spatial separation distance. The spatial correlation function and semi-variogram convey the same information, as is shown in the following equation:

$$\rho(d_{ij}) = 1 - \frac{\gamma(d_{ij})}{\mathrm{Var}(X)} \tag{3}$$

where $\rho(d_{ij})$ is the correlation coefficient for points separated by distance d_{ij} and Var(X) is the variance of the variable X. The sample correlation coefficients can be computed by the following equation:

$$\rho(d_{ij}) = \frac{(1/n)\sum_{1}^{n}\left[X_i - \bar{X}_i\right]\left[X_j - \bar{X}_j\right]}{\sigma_x^2} \tag{4}$$

In all of the above equations i and j are two stations separated by a distance d_{ij}, n is the number of pairs of observations, X is the sample value, $\bar{X}$ is the mean, and σ_X^2 is the variance of the individual weather variable.

5.3. Cross-semi-variogram

The co-regionalization of two variables Z_1 and Z_2 (describing spatial correlation of two different weather variables) is summarized by a sample cross-semi-variogram:

$$\gamma_{12}(d_{ij}) = \frac{1}{2n}\sum_{1}^{n}\left[Z_1(i) - Z_1(j)\right]\left[Z_2(i) - Z_2(j)\right] \tag{5}$$

where n is the number of pairs of observation separated by a distance d_{ij}, and i and j are two sampling points. Variables Z_1 and Z_2 do not necessarily need to have the same number of samples, but the cross-semi-variogram is calculated using only the locations where both variables are measured. Unlike semi-variograms, cross-semi-variograms can be negative, if the relationship between Z_1 and Z_2 is negative. The cross-semi-variogram is used to enhance kriging estimates via co-kriging.

Spatial correlation structures, as described by the semi-variogram, cross-semi-variogram, or spatial correlation function, may be anisotropic, i.e. may vary with direction.

5.4. Interpolation methods

There are a number of commonly used interpolation techniques described in the literature, such as simple average, Thiessen polygon, classical polynomial interpolation, inverse distance, multi quadric interpolation, optimal interpolation, kriging and others. In this study two of the most promising interpolation techniques are applied for designing weather monitoring networks: inverse distance interpolation was chosen for its simplicity, and kriging interpolation and its extension co-kriging were chosen to take advantage of the spatial correlation structure described by the semi-variogram and cross-semi-variograms.

5.4.1. Inverse distance interpolation

As is obvious from the name of the interpolation technique, the weighting factor is inversely proportional to the distance. The weights of this interpolation technique are solely a function of the distance between the point of interest, with coordinates (H_0,V_0), and the sampling points (H_i,V_i) for $I=1,2,\cdots,n$. Considering the distance d_{0i} between these two points, the weight of a sampling point (H_i,V_i) is in the form

$$w_i = \frac{f(d_{0i})}{\sum_{j=1}^{n} f(d_{0j})} \tag{6}$$

where $f(d_{0i})$ represents a given function of the distance d_{0i}. The common form of this function is

$$f(d_{0i}) = \frac{1}{(d_{0i})^{b_2}} \tag{7}$$

where b_2 is an appropriate constant. The weight w_1 approaches zero as the distance d increases. When the parameter b takes the value of one or two, the technique is called inverse distance interpolation or inverse squared distance interpolation, respectively. Ripley (1981) recommended the use of one, two and four as the values for b_2. The inverse distance technique does not take advantage of the spatial correlation structure explicitly. However, as we have already noted that these correlation structures tend to be linear, we might guess that inverse distance weighting would work fairly well.

5.4.2. Kriging

The kriging technique is somewhat more sophisticated than those mentioned above, and the weights are directly related to the spatial correlation structure (semi-variogram). There are two basic advantages to kriging over simpler interpolation techniques. First, kriging provides the best linear unbiased estimator (BLUE) to the interpolation problem, given the basic assumptions of no trend and a model for the semi-variogram. Both of these assumptions apply in the present study as Eq. (2) is used to calculate the semi-variogram, and good models are obtained using a linear fit for the semi-variogram. Second, an analysis of interpolation error is provided (Tabios and Salas, 1985). (For a detailed description and the derivation of the kriging system of equations, see Clark (1982) and Tabios and Salas (1985).)

5.4.3. Co-kriging

Co-kriging extends the principle of optimal estimation using regionalized variable theory from that of a single property to situations where there are two or more co-regionalized properties. In other words co-kriging takes advantages of inter-variable correlation. Co-kriging is more efficiently used where one variable may not have been sampled sufficiently (owing to experimental difficulties, high costs, etc.) to provide estimates of acceptable precision. Estimation precision can be improved by utilizing the spatial correlation between the under-sampled (primary) variable and other, more frequently sampled co-variables. The concepts of co-kriging discussed here assume only one co-variable, but the equations are readily expanded to include additional co-variables.

The co-kriged value of the under-sampled, or primary, variable, Z_2, is computed as a weighted average of the observed values of the co-variable, Z_1, and Z_2 occurring in the estimation neighborhood of each kriged point. The co-kriged value $\hat{Z}_2$ at point zero is

$$\hat{Z}_2(0) = \sum_{i=1}^{n_1} W_{1i} Z_1(i) + \sum_{j=1}^{n_2} W_{2j} Z_2(j) \tag{8}$$

where W_{1i} and W_{2j} are the weights associated with Z_1 and Z_2, respectively, and n_1 and n_2 are the number of neighbors of Z_1 and Z_2 involved in estimating $\hat{Z}_2$ at point of interest zero, respectively.

The weights on observed values of Z_1 and Z_2 are chosen so that the estimate is unbiased with minimum variance, just as in kriging. However, the solution of the co-kriging system of equations for the weights is obtained using the semi-variograms and cross-semi-variograms of each Z_1 and Z_2 with the kriged location zero. The resulting system of equations is more complex than in kriging. Like kriging, solution of the co-kriging system also yields the co-kriging estimation variance for interpolated location zero. The equations of co-kriging system have been presented in full by Isaaks and Srivastava (1989) and by Vieira (1983). The co-kriging system requires at least one sample point of both the primary variable and co-variable properties within the estimation neighborhood. If the primary variables and co-variables are measured at all sampling sites in the neighborhood, then co-kriging yields the same estimate as kriging of the primary variable alone. In this case, co-kriging is of no value.

6. Evaluating network performance

To evaluate the performance of a given interpolation method or a network configuration, we need a criterion for measuring performance. One such criterion is root mean squared interpolation error (RMSIE), defined below. The RMSIE is widely used when one is most interested in the few largest errors as opposed to the simple average of all errors. The RMSIE is similar in form to the sample standard deviation and has the same form as kriging error:

$$\text{RMSIE} = \left[\frac{\sum_{i=1}^{n} \left[\text{actual } Et_r - Et_r(\text{estimated})\right]^2}{n} \right]^{0.5} \tag{9}$$

A key question that arises is whether it is better to interpolate individual weather variables or to interpolate computed Et_r. To answer this we shall use three different estimates of Et_r at a particular station location. The first and most accurate is the actual Et_r estimate computed from the observed data at that particular station. The second, which we shall call Et_r (computed) is estimated from the interpolated values of weather variables from surrounding stations to that particular station (and ignoring data from the station itself). The third, which we shall call Et_r (interpolated) is estimated at that particular station from the interpolation of computed Et_r values at surrounding stations (again ignoring the station's own data). Et_r (computed) and Et_r (interpolated) are compared with the actual Et_r estimate using the RMSIE. The interpolation location for which actual Et_r estimates are available is sometimes called a 'fictitious point'.

7. Results and discussions

7.1. Spatial structures of weather variables

Calculated spatial structures for daily summer (21 June–21 September) data of 1989 are given in Table 2. All of the spatial functions may be represented by a linear model as

$$Xi = A + B(d_{ij}) \tag{10}$$

where Xi is an individual weather variable, d_{ij} is the distance (in km), and A and B are the coefficients of the equation, computed by the regression analysis. The values of A, B and correlation coefficient r^2 are given in Table 2. All these equations are based on the spatial structure after mean (drift or trend) is removed. Both the present study and that by Harcum and Loftis (1987a) investigated whether spatial correlation was a function of direction for the variables studied in the High Plains regions of Colorado, Kansas, and Nebraska, and found that isotropy was a reasonable assumption. If this assumption were violated, the kriging estimates would not be statistically optimal but still might be shown to be very good.

The linear function of the semi-variogram increases apparently without bound, implying that the function does not reach its sill (a constant value for long lag distances indicating no correlation) for the case study (i.e. weather variables are spatially correlated out to a distance of at least 600 km). The spatial structure persists for such a long distance because general weather patterns affect large areas. Also, temporal patterns are not considered in our model other than our consideration of summer data only. We recognize of course that stations which are close together have similar seasonal patterns. If we had a very good seasonal model for each location, we could remove the daily seasonal component and reduce the apparent spatial correlation (Harcum and Loftis, 1987a, Harcum and Loftis, 1987b). As we mentioned above, however, long-term data would be needed for this approach. Although such records are available for many locations in the USA, this might not be true for the developing countries which are a focus of this research. Furthermore, the modeling and estimation approach used here considering spa-

Table 2
Spatial structures of weather variables of daily data ($Xi = A + B \times d_{ij}$), and cross-semi-variograms of weather variables of daily data ($WVi = A + B \times d_{ij}$), both for summer 1989 (92 days)

Spatial structure	A	B	r^2
Weather variable (Ξ)			
Maximum temperature (°C)			
Semi-variogram	−0.590	0.025295	0.88
Correlation coefficient	1.005	−0.000610	0.90
Minimum temperature (°C)			
Semi-variogram	1.509	0.008143	0.43
Correlation coefficient	0.940	−0.000270	0.49
Mean temperature (°C)			
Semi-variogram	−0.170	0.012511	0.76
Correlation coefficient	0.993	−0.000400	0.80
Relative humidity (%)			
Semi-variogram	11.940	0.312067	0.61
Correlation coefficient	0.971	−0.001040	0.72
Wind run (km day^{-1})			
Semi-variogram	640.300	14.751990	0.63
Correlation coefficient	0.965	-0.001290	0.68
Solar radiation (MJ m^2 day^{-1})			
Semi-variogram	0.769	0.055201	0.87
Correlation coefficient	0.975	−0.001320	0.89
Et_r (mm day^{-1})			
Semi-variogram	0.138	0.008577	0.76
Correlation coefficient	0.990	−0.000900	0.86
Maximum temperature × solar radiation			
Cross-semi-variogram	−0.986	0.022100	0.81
Cross-semi-variogram	**A**	**B**	**r^2**
Weather variable (WVi)			
Maximum temperature × solar radiation	−0.986	0.0221	0.81
Maximum temperature × solar radiation, summer 1990	−1.614	0.0280	0.81
Maximum temperature × Et_r	−0.431	0.012	0.85
Mean temperature × Et_r	−0.357	0.0078	0.74
Relative humidity × Et_r	−0.470	−0.0444	0.69
Wind run × Et_r	10.602	0.1261	0.39
Solar radiation × Et_r	−0.0532	0.0157	0.84

tial correlation is much simpler than that of Harcum and Loftis (1987a, Harcum and Loftis, 1987b).

As mentioned above, all of the variables studied show strong spatial correlation out to long lag distances. For some variables, however, the linear model prediction of the correlation function or semi-variogram is less precise than for others. Table 2 suggests that solar radiation, Et_r, mean temperature and maximum temperature are 'better correlated' in space (in that they have less scatter in their correlation functions) than wind run, relative humidity and minimum temperature, indicating that the latter three elements are more variable in space; that is, they are more affected by local features of the landscape or localized weather patterns. The variables with the noisier spatial correlation are included in our Class B weather

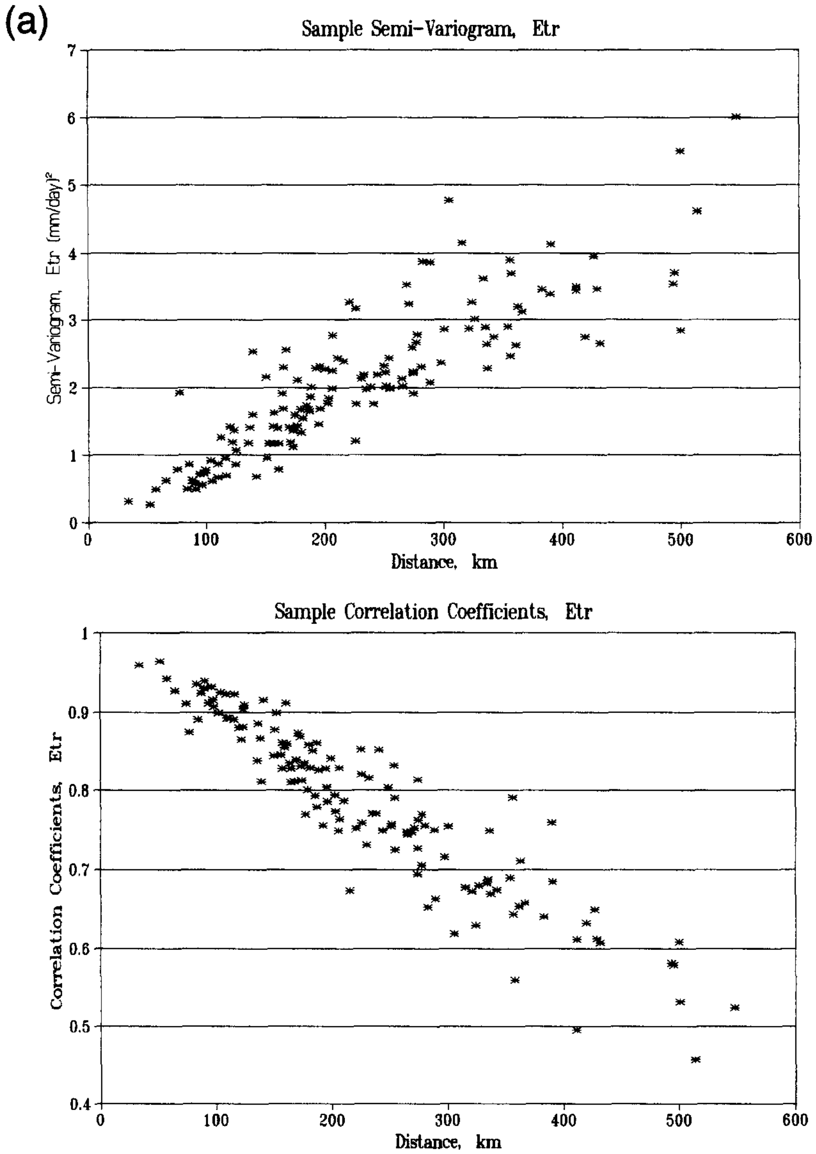


Fig. 2. (a) Spatial structures of Et_r (computed) for daily data of summer 1989. (b) Sample cross-semi-variogram of solar radiation and Et_r and of relative humidity and Et_r for summer 1989 data.

stations. Examples of the spatial correlation function and semi-variogram are shown in Fig. 2(a), and examples of cross-semi-variograms are shown in Fig. 2(b).

7.2. Application of interpolation methods

For the case study data, the Et_r (computed) has lower RMSIE values than those of Et_r (interpolated) in general. Fig. 3 shows the comparison between RMSIE of these two estimates of Et_r using Dickens and Garden City as fictitious points, one by one, when 11 Class A weather stations were used for interpolation of summer 1989 data. The RMSIE values were lower for estimation of Et_r by interpolating weather variables and computing Et_r than for estimation of Et_r by interpolation of the computed Et_r values, but the difference is small. The significant point is that one may interpolate individual weather variables and then calculate Et_r effectively,

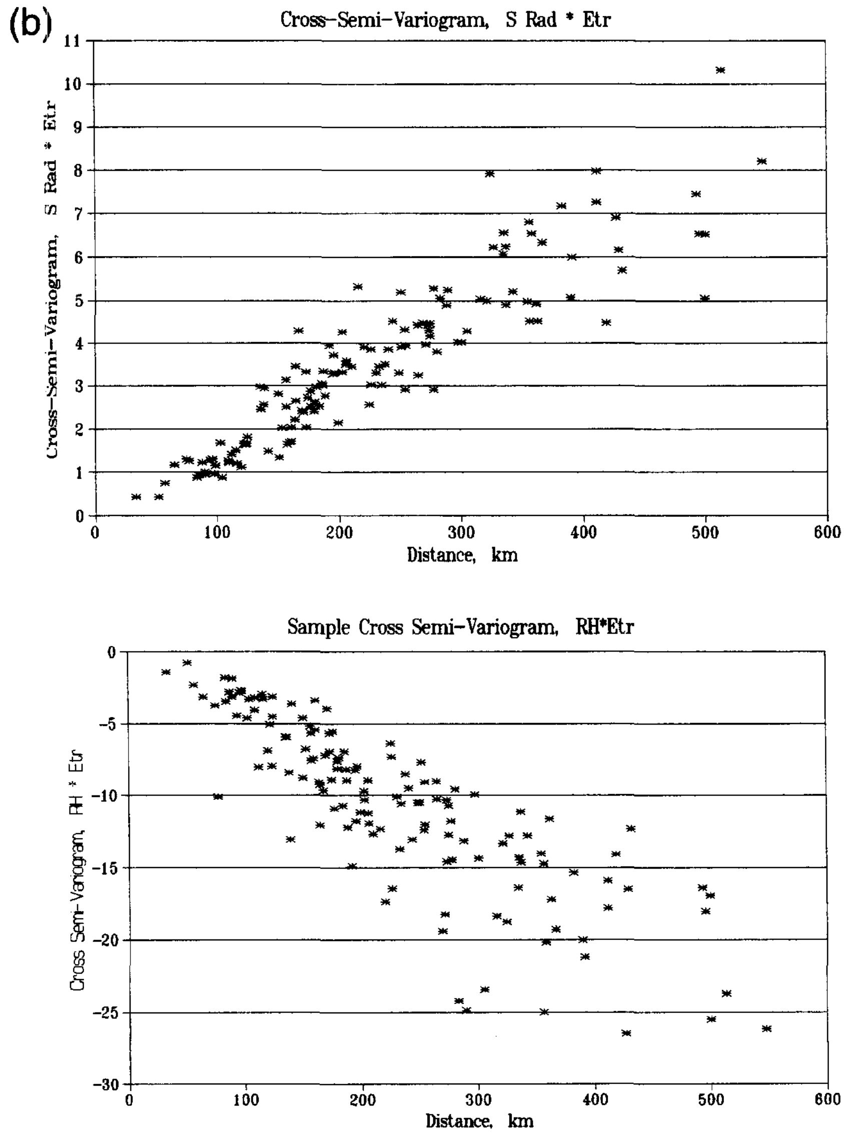


Fig. 2. (continued).

which is necessary when applying partial weather station networks.

Tables 3 and 4 provide a comparison of the three interpolation methods for a variety of network configurations (to be discussed below) using Dickens and Garden City as fictitious points. Several time averaging periods are shown as well. If one were using a center pivot to irrigate daily, then daily values of weather variables and Et_r might be most appropriate. For less frequent irrigation, moving averages might be more appropriate. Kriging emerges as a clear winner in this comparison. An example of

Table 3
Performance of alternative weather station network configurations and interpolation methods in estimating Et_r at Dickens, Nebraska; table values show reduction in RMSIE compared with the first configuration of six Class A weather stations; additional improvement in RMSIE of Et_r (mm day^{-1}) at Dickens for daily summer 1989 data by using co-kriging to 'convert' Class B stations into Class A stations is also shown

Data used	Number of stations	RMSIE (mm)			% Reduction in RMSIE		
		Inv.	Inv.^{2}	Kriging	Inv.	Inv.^{2}	Kriging
Daily, summer 1989	6A	1.4	1.4	1.3	0	0	0
	6A + 5B	1.1	1.0	0.8	19	28	36
	6A + 10B	1.0	0.9	0.7	30	35	44
	11A	1.0	0.8	0.6	28	40	53
	11A + 5B	0.9	0.8	0.5	37	44	58
	16A	0.9	0.8	0.6	38	41	52
3 day averages, summer 1989	6A	1.1	1.1	1.0	0	0	0
	6A + 5B	0.8	0.7	0.6	23	35	45
	6A + 10B	0.7	0.7	0.5	32	38	51
	11A	0.8	0.6	0.4	31	45	62
	11A + 5B	0.7	0.6	0.4	39	45	61
	16A	0.7	0.6	0.5	39	41	54
3 day moving averages, summer 1989	6A	1.1	1.1	1.0	0	0	0
	6A + 5B	0.8	0.7	0.6	23	35	44
	6A + 10B	0.7	0.7	0.5	32	38	51
	11A	0.8	0.6	0.4	28	43	57
	11A + 5B	0.7	0.6	0.4	36	43	60
	16A	0.7	0.7	0.5	35	38	54
7 day averages, summer 1989	6A	0.8	0.8	0.7	0	0	0
	6A + 5B	0.6	0.5	0.4	26	41	45
	6A + 10B	0.5	0.5	0.4	31	37	50
	11A	0.6	0.4	0.3	28	45	55
	11A + 5B	0.5	0.5	0.3	33	38	56
	16A	0.6	0.6	0.4	28	29	47
7 day moving averages, summer 1989	6A	0.9	0.9	0.8	0	0	0
	6A + 5B	0.7	0.5	0.4	24	40	49
	6A + 10B	0.6	0.6	0.4	29	35	54
	11A	0.6	0.5	0.3	27	44	59
	11A + 5B	0.6	0.6	0.3	31	37	60
	16A	0.6	0.6	0.4	27	30	52

Number of stations	Additional improvement					
	RMSIE (mm)			% Reduction in RMSIE from original configuration		
	Inv.	Inv.^{2}	Kriging	Inv.	Inv.^{2}	Kriging
6A + 5B	1.1	1.0	0.8	2	2	2
6A + 10B	0.9	0.8	0.6	7	6	11
11A + 5B	0.8	0.8	0.5	4	0	−2

Inv., Inverse distance interpolation; Inv.^{2}, inverse squared distance interpolation.

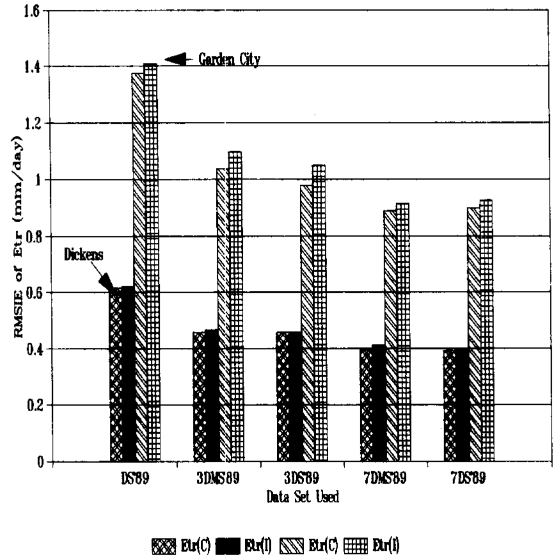


Fig. 3. Comparison of RMSIE of Et_r (computed) and Et_r (interpolated) at Dickens and Garden City for different data sets. DS'89, Daily data for summer 1989; 3DS'89, 3 day average data for summer 1989; 3DMS'89, 3 day moving averages for summer 1989; 7DS'89, weekly averages for summer 1989; 7DMS'89, weekly moving averages for summer 1989.

co-kriging is also included for one of the data sets used in this study. Table 3 shows the additional improvement when co-kriging was applied to the mixture of Class A and Class B network configurations.

In these configurations all Class B weather stations were 'converted' to Class A stations, one by one (by 'converted' we mean that solar radiation was estimated by spatial interpolation of the Class B weather stations, using co-kriging) and then the estimated values at these stations were interpolated to Dickens. These new RMSIE values are listed in Table 3. In this co-kriging 'conversion' process, the cross-semi-variogram of maximum temperature and solar radiation was used. This table also shows the per cent reductions in RMSIE values from the original configurations. The advantage of co-kriging in this example is up to 10% reduction in RMSIE compared with ordinary kriging. The advantage disappears in the most dense network configuration. The inverse distance methods were also successful in 'converting' Class B stations to Class A but were less effective than co-kriging.

The RMSIE of individual weather variables can also be computed. If measurement of individual

Table 4
Improvement in RMSIE of Et_r (mm day^{-1}) at Garden City for different data sets using different classes of weather stations, based on reference configuration

Data used	Number of stations	RMSIE (mm)			% Reduction in RMSIE		
		Inv.	Inv.^{2}	Kriging	Inv.	Inv.^{2}	Kriging
Daily, summer 1989	6A	1.8	1.5	1.4	0	0	0
	6A + 5B	1.9	1.6	1.4	−1	−4	2
	6A + 10B	1.9	1.7	1.4	−5	−10	1
	11A	1.9	1.7	1.4	−3	−7	2
	11A + 5B	2.0	1.8	1.4	−7	−13	2
	16A	2.0	1.8	1.4	−10	−16	2
3 day moving averages, summer 1989	6A	1.4	1.2	1.1	0	0	0
	6A + 5B	1.4	1.2	1.1	−1	−5	2
	6A + 10B	1.5	1.3	1.1	−5	−12	1
	11A	1.4	1.3	1.1	−3	−8	4
	11A + 5B	1.5	1.3	1.1	−7	−14	3
	16A	1.5	1.4	1.0	−9	−18	6
7 day moving averages, summer 1989	6A	1.1	0.9	1.0	0	0	0
	6A + 5B	1.1	0.9	0.9	2	−3	4
	6A + 10B	1.1	1.0	0.9	−1	−8	3
	11A	1.1	0.9	0.9	1	−5	5
	11A + 5B	1.1	1.0	0.9	−2	−10	4
	16A	1.1	1.0	0.9	−5	−13	7

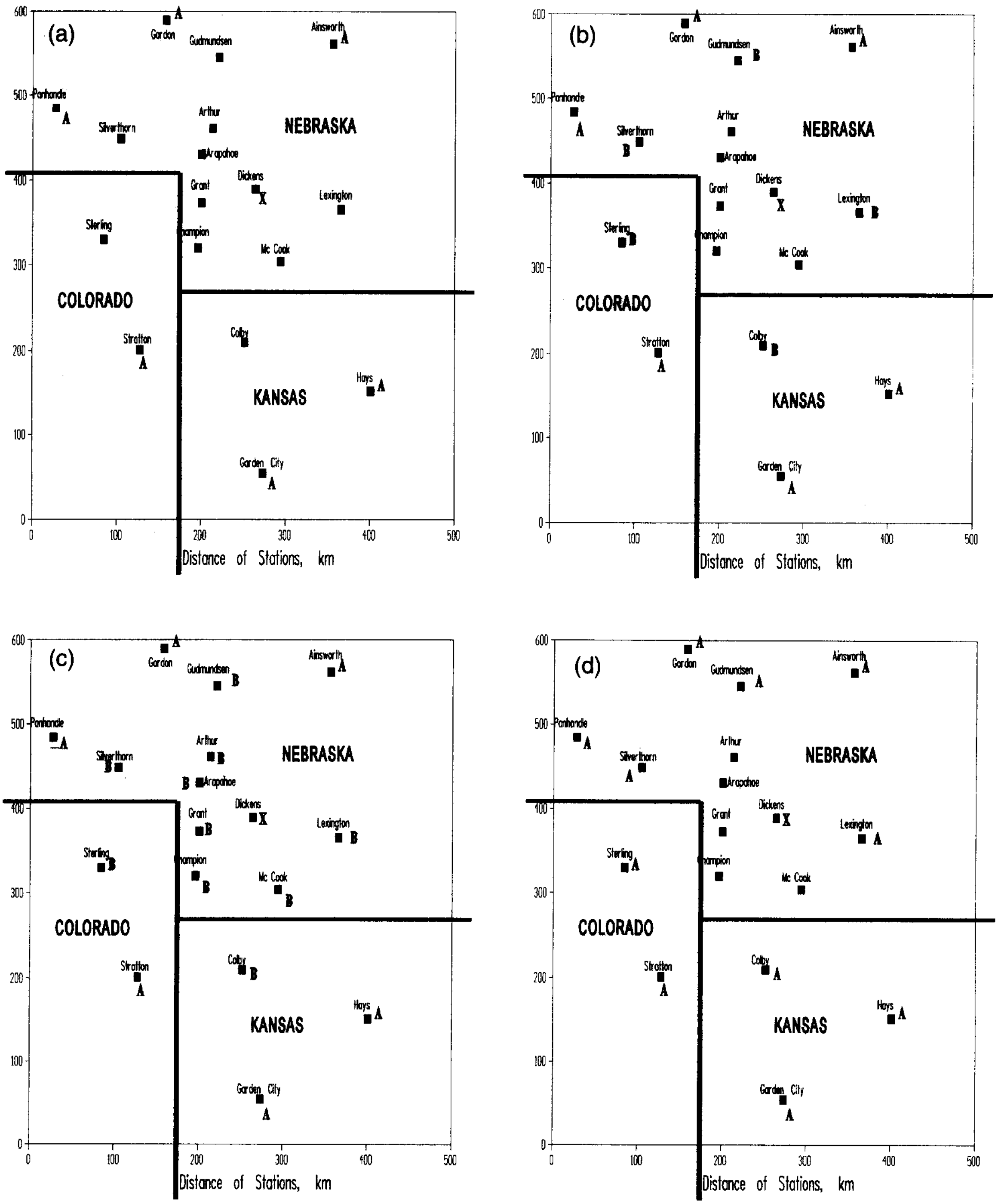


Fig. 4. Selected case-study network configurations. (a) Six Class A weather stations; (b) six Class A and five Class B weather stations; (c) six Class A and ten Class B weather stations; (d) 11 Class A weather stations; (e) 11 Class A and five Class B weather stations.

weather variables were a major objective, a network design could be based on these errors rather than errors in Et_r.

7.3. Partial weather station network evaluation

The results of RMSIE values using full weather stations were compared with those obtained with the introduction of Class B weather stations in the case study area. These RMSIE values are reduced when Class B stations are added to the network. The advantage of partial weather stations is of course cost. They may be used to reduce interpolation errors when inserted into a network of Class A weather stations or can be a part of the design of a new network.

A case study was used to evaluate performance of 'mixed' networks by using the Dickens and Garden City locations as fictitious points. Six network configurations were used in this study, some showing mixture of Class A and Class B weather stations. One such configuration is shown in Fig. 1, in which all 16 stations are used as Class A stations. Five other network configurations are shown in Fig. 4.

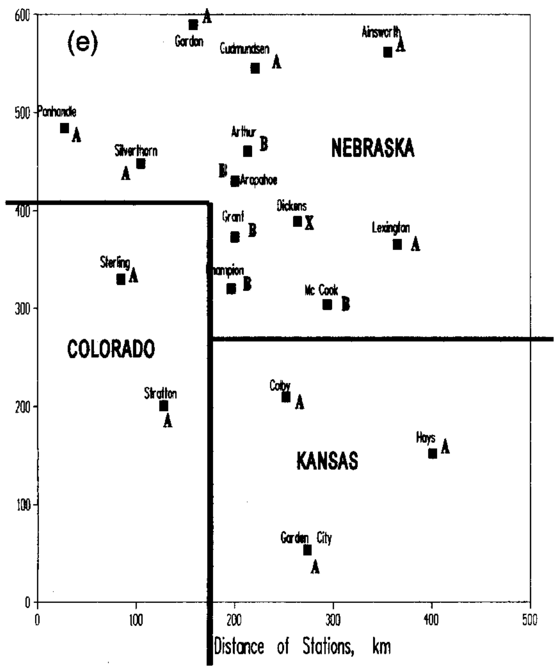


Fig. 4. (continued).

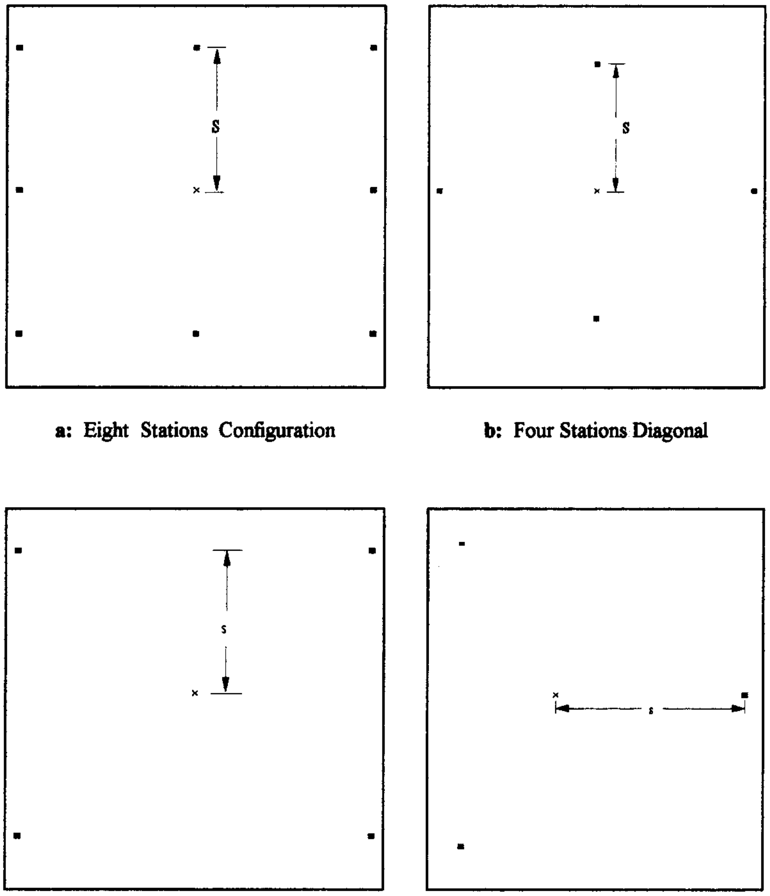
Table 3 shows the per cent reduction (from a reference configuration, which is six Class A weather stations) in RMSIE values of Et_r at Dickens by using different numbers of Class A and Class B weather stations. Both Class A and Class B weather stations are actually full weather stations in this study; we simply ignore the available solar radiation at the assumed Class B weather stations.

The reduction in RMSIE values compared with the first configuration ranges from 35 to 55% for the chosen configurations and data sets used for interpolation at Dickens, a station in the middle of the network. Table 4 presents the same results for Garden City, one of the outlying fringe stations in the network. In this case, the reduction in RMSIE was much less, ranging from 1 to 5%. Tables 3 and 4 suggest that the improvement in estimation obtained by adding five Class B stations is almost as great as that of five Class A stations.

8. Network design methods

One rarely has the flexibility to locate stations ideally. However, it is instructive to consider the question of how one would arrange stations in an ideal case. Four configurations for the general network design were compared in terms of kriging variances for an assumed spatial structure. Fig. 5 shows these configurations: (a) eight-station configuration; (b) four-station-diagonal configuration; (c) four-station-square configuration; (d) three-station configuration. The results are shown in Fig. 6. These results show that the kriging variances for eight-station and four-station-diagonal configurations are equal and are lower than the kriging variances of the other two configurations. Interestingly, the four corner stations of the eight-station configuration have very small negative kriging weights for interpolation at the point of interest and add nothing to the estimation. Thus a four-station diagonal arrangement would be preferred in an ideal case.

As a convenient tool for comparison of alternative configurations, the RMSIE can be plotted against the number of weather stations and their cost. Fig. 7 shows such graphs. In this figure, Class A weather stations are assumed to have a cost of \$10000, and that of Class B weather stations is \$4500. Using



a: Eight Stations Configuration

b: Four Stations Diagonal

c: Four Stations Square

d: Three Stations Configuration

Fig. 5. Different general configurations of network design. ■, Weather station; ×, point of interest.

graphs such as this, one can easily see the tradeoffs between interpolation error and budget, and can choose the number and type of weather stations for redesign of this network. If no historic data are available for estimating spatial correlation structures, then data from a similar climatic region may be used. It should be recalled, however, that the method proposed here does not require a long-term record.

9. Conclusions

1. The spatial structures of all weather variables used to compute Et_r are found to follow linear models up to a distance of 600 km in the case study. Spatial correlations of maximum temperature, solar radiation, Et_r and mean temperature are better than those of relative humidity, wind run and minimum temperature. Therefore it is logical to include the latter three variables in partial (Class B) weather stations at closer spacings than full (Class A) weather stations.
2. Of the three interpolation techniques studied, kriging interpolation provided the lowest RMSIE, and inverse distance-square interpolation provided a lower RMSIE than inverse distance interpolation. Co-kriging was found to improve estimates over ordinary kriging when partial weather stations are included in an agricultural meteorology network.

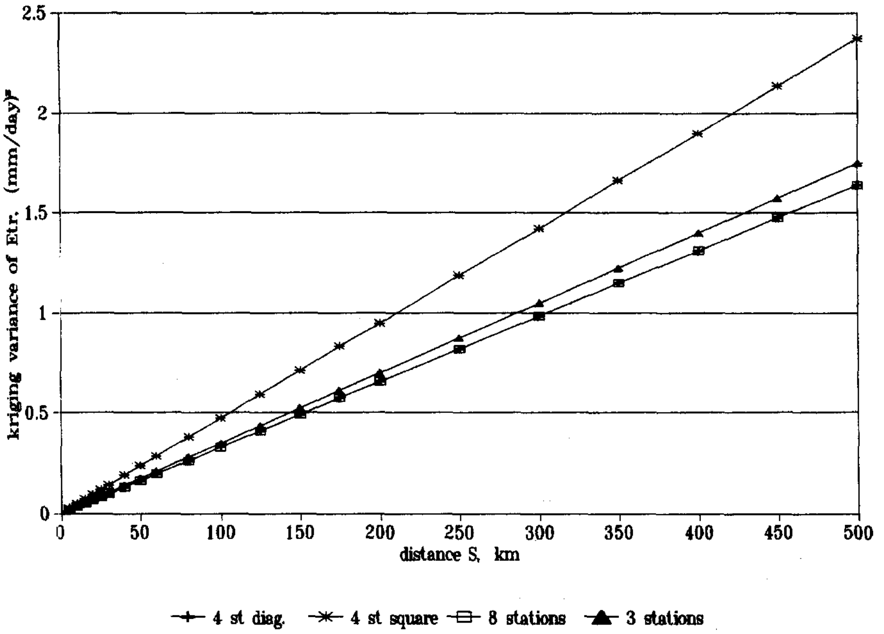


Fig. 6. Comparison of kriging variances for general configurations of different network design.

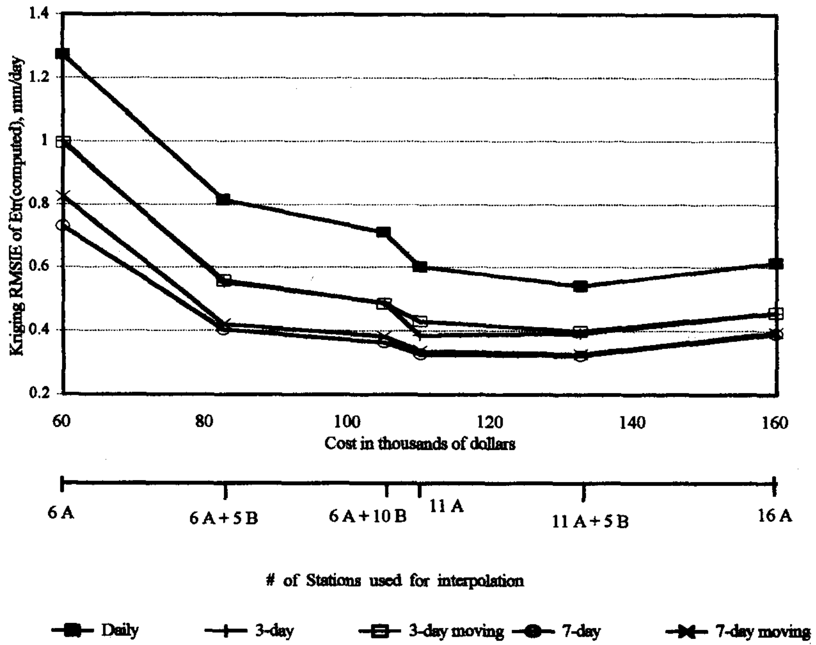


Fig. 7. Cost effectiveness of different classes of weather stations on kriging RMSIE at Dickens for summer 1989 data (Class A weather station costs US$10 000; Class B weather station costs US$4500).

3. The kriging method suggested here does not use temporal correlation of weather variables and therefore does not require a long-term record of weather data.
4. For the case study, estimates of Et_r from interpolating weather variables and computing Et_r were comparable with those from computing Et_r and then interpolating. Thus, the interpolation of weather variables required when using partial weather station networks results in little loss of accuracy.
5. A mixture of full and partial weather stations may have a distinct cost advantage compared with a network composed of full weather stations only.
6. A comparison of kriging errors and costs of alternative weather monitoring network configurations may be used a basis for network design or improvement. This design approach may be most applicable to the Third World, where resources are severely limited, but it also can be used in developed countries to reduce network costs.

Appendix A

A.1. Computation of reference evapotranspiration (Et_r)

The Penman combination equation, modified for estimating alfalfa-based reference Et_r (in mm day^{-1}) has been given by Jensen et al. (1990):

$$LEt_r = \frac{\Delta}{\Delta+\gamma}(\boldsymbol{R}_n - \boldsymbol{G}) + \frac{\gamma}{\Delta+\gamma}6.43W_f(e_a - e_d) \qquad \text{(A1)}$$

where L is latent heat of vaporization (MJ kg^{-1}) and can be computed as

$$L = 2.501 - 2.361\times10^{-3}T \quad (T\,°\mathrm{C}) \qquad \text{(A2)}$$

Et_r is reference evapotranspiration (mm day^{-1}), Δ is slope of the vapor pressure–temperature curve (kPa °C^{-1}), γ is the psychrometric constant (kPa °C^{-1}), $\boldsymbol{R}_n$ is net radiation (MJ m^{-2} day^{-1}), $\boldsymbol{G}$ is soil heat flux to the surface (MJ m^{-2} day^{-1}) (where

$$\boldsymbol{G} = 0.377[T - (T_{-1} + T_{-2} + T_{-3})/3] \qquad \text{(A3)}$$

where T is the mean daily air temperature (°C) and T_{-i} is the mean air temperature for the ith previous day; the daily values of $\boldsymbol{G}$ are small and assumed to be zero in this study), W_f is wind function, $e_a - e_d$ is mean daily vapor pressure deficit (kPa), and 6.43 is a constant of proportionality (MJ m^{-2} day^{-1} kPa^{-1}).

Bosen's formula for saturation vapor pressure, when differentiated, gives Δ that varies with temperature (Jensen et al., 1990):

$$\Delta = 0.200(0.00738T + 0.8072)^7 - 0.000116 \qquad \text{(A4)}$$

and γ can be computed as

$$\gamma = \frac{c_p P}{0.622L} \qquad \text{(A5)}$$

where P is the average station barometric pressure (kPa) (where

$$P = 101.3 - 0.01055E \qquad \text{(A6)}$$

and E is elevation above mean sea level (m)), 0.622 is the ratio of the molecular mass of water to the apparent molecular mass of dry air, and c_p is specific heat of air at constant pressure ($c_p = 1.003\times10^{-3}$ MJ kg^{-1} °C^{-1}).

$\boldsymbol{R}_n$ can be estimated by the solar radiation data:

$$\boldsymbol{R}_n = (1-\alpha)\boldsymbol{R}_s - \boldsymbol{R}_b \qquad \text{(A7)}$$

where α is reflected shortwave radiation fraction, called albedo ($\alpha = 0.23$ for commercial irrigated crops), $\boldsymbol{R}_s$ is incoming shortwave solar radiation (MJ m^{-2} day^{-1}), and $\boldsymbol{R}_b$ is net outgoing radiation (MJ m^{-2} day^{-1}), which can be computed as

$$\boldsymbol{R}_b = \left(a\frac{\boldsymbol{R}_s}{\boldsymbol{R}_{so}} + b\right)\boldsymbol{R}_{bo} \qquad \text{(A8)}$$

where $\boldsymbol{R}_{bo}$ is net outgoing longwave radiation on a clear day and may be estimated as follows:

$$\boldsymbol{R}_{bo} = \left(a_1 + b_1\sqrt{e_d}\right)(4.903\times10^{-9})T_k^4 \qquad \text{(A9)}$$

where T_k is average daily air temperature (K). The values of experimentally determined coefficients a, b, a_1 and b_1 are taken from Jensen et al. (1990).

R_{so} is solar radiation for cloudless skies ($\mathrm{MJ\,m^{-2}\,day^{-1}}$) and can be computed as follows (Jensen et al., 1990):

$$R_{so} = A' + B'\cos[(2\pi D/365) - C'] \tag{A10}$$

where $A' = 31.55 - 0.273\mathrm{Lat} + 0.0008E$, $B' = -0.299 + 0.268\mathrm{Lat} + 0.0004E$, D is the calendar day (1–365), C' is the phase constant for the longest day (normally 172), and Lat is latitude (°N).

The W_f is usually determined by regression analysis:

$$W_f = a_w + b_w U_2 \tag{A11}$$

where a_w and b_w are the regression coefficients, and U_2 is the daily wind travel at 2 m above the ground ($\mathrm{km\,day^{-1}}$). The equations for the regression coefficients have been described by Jensen et al. (1990):

$$a_w = 0.4 + 1.4\exp\left\{-\left[\frac{(D-173)}{58}\right]^2\right\} \tag{A12}$$

and

$$b_w = 0.007 + 0.004\exp\left\{-\left[\frac{(D-243)}{80}\right]^2\right\} \tag{A13}$$

where D in the above equations is the calendar day of the year. Jensen et al. (1990) also recommended the values of these coefficients for the alfalfa-related combination equations for humid and arid areas.

The vapor pressure deficit ($e_a - e_d$) was calculated using the following expression for estimating the saturation vapor pressure given by Bosen (Jensen et al., 1990) and using observed values of relative humidity to calculate actual vapor pressure:

$$e_a = 3.38639\left[(0.00738T + 0.8072)^8 - 0.000019|1.8T + 48| + 0.001316\right] \tag{A14}$$

References

Amegee, K.Y., 1985. Application of geostatistics to regional evapotranspiration. Ph.D. Dissertation, Oregon State University, Corvallis, 176 pp.

Clark, I., 1982. Practical Geostatistics. Elsevier Applied Science, London, 129 pp.

Duke, H.R., Buchleiter, G.W. and Heerman, D.F., 1985. The meteorologic and agronomic theory incorporated in the USDA ET computation program SCHED. USDA-ARS, Colorado State University, Fort Collins, 11 pp.

Gandin, L.S., 1970. The planning of meteorological station networks. WMO Tech. Note, 111.

Harcum, J.B. and Loftis, J.C., 1987a. Spatial interpolation of Penman evapotranspiration. Trans. ASAE, 30(1): 129–136.

Harcum, J.B. and Loftis, J.C., 1987b. Design and evaluation of regional weather monitoring networks. Trans. ASAE, 30(6): 1673–1678.

Hubbard, K.G. 1994. Spatial variability of daily weather variables in the high plains of the USA. Agric. For. Meteorol., 68: 29–41.

Hubbard, K.G., Rosenberg, N.J. and Nielsen, D.C., 1983. Automated weather data network for agriculture. J. Water Resour. Plann. Manage., 109: 213–222.

Isaaks, E.H. and Mohan Srivastava, R., 1989. Applied Geostatistics. Oxford University Press, New York, 561 pp.

Jensen, M.E. (Editor), 1973. Consumptive use of water and irrigation water requirements. Irrig. Drainage Div., American Society of Civil Engineers, New York, 215 pp.

Jensen, M.E., Burman, R.D. and Allen, R.G., 1990. Evapotranspiration and Irrigation Water Requirements. A manual prepared by the Committee on Irrigation Water Requirements of the Irrigation and Drainage Division of the American Society of Civil Engineers, New York, 332 pp.

Ripley, B.D., 1981. Spatial Statistics. Wiley, New York, 252 pp.

Tabios, G.Q. and Salas, J.D., 1985. A comparative analysis of techniques for spatial interpolation of precipitation. Water Resour. Bull., 21(3): 365–380.

Vieira, S.R., 1983. Geostatistical analyses of some agronomical observations. Ph.D. Dissertation, University of California, Davis, 261 pp.