JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION
VOL. 37, NO. 3 AMERICAN WATER RESOURCES ASSOCIATION JUNE 2001

STRATIFICATION OF VARIABILITY IN RUNOFF AND SEDIMENT YIELD BASED ON VEGETATION CHARACTERISTICS[1]

Daryl E. Mergen, M. J. Trlica, James L. Smith, and Wilbert H. Blackburn[2]

ABSTRACT: Runoff and sediment yield were collected from 100 plots during simulated rainfalls (100 mm/hr for 15 minutes) at antecedent soil moisture conditions. A clustering technique was used to stratify the variability of a single data set within a sagebrush-grass community into four groups based on vegetation life form and amount of cover. The four cluster groups were grass, grass/shrub, shrub, and forb/grass and were found to be significantly different in plant height, surface roughness, soil bulk density, and soil organic matter. Stepwise multiple regression analyses were performed on the single data set and each cluster group. Results for individual groups resulted in more robust predictive equations for runoff (r^2 = 0.65-0.73) and sediment yield (r^2 = 0.37-0.91) than for equations developed from the single data set (r^2 = 0.56 for runoff and r^2 = 0.27 for sediment yield). The standard errors of the cluster group regression equations were also improved in three of the four group equations for both runoff and sediment yield compared to the single data set. Runoff was found to be significantly less ($p < 0.01$) in the forb/grass group compared with other vegetation cluster groups, but this was influenced by four plots that produced little or no runoff. Sediment yield was not found to be significantly different among any cluster groups. Discriminant analysis was then used to identify important variables and develop a model to classify plots into one of the four cluster groups. The discriminant model could be incorporated into rangeland hydrology and erosion models. The percentage cover of grasses, shrubs, litter, and bare ground effectively stratified about 12 percent of the variation observed in runoff and 26 percent of the variability for sediment yield as determined by r^2.
(KEY TERMS: spatial variability; cluster analysis; discriminant analysis; erosion; infiltration.)

INTRODUCTION

Considerable spatial variability in hydrologic and erosion processes exists on semiarid rangelands as a result of plant and soil differences. Vegetation has been shown to influence macro- and microtopography and soil characteristics (Blackburn, 1975; Thurow *et al.,* 1988). This causes great spatial variability observed in runoff, infiltration, and erosion (Devaurs and Gifford, 1984; Simanton *et al.,* 1991; Pierson *et al.,* 1994). Blackburn *et al.* (1992) concluded that plant form affected the variability of soil surface properties, and these properties in turn affected the variability in hydrologic and erosion processes (Pierson *et al.,* 1994). Hydrological variability caused by vegetation and soil differences make accurate prediction of runoff, infiltration, and sediment yield difficult (Blackburn *et al.,* 1992, Blackburn and Pierson, 1994).

Considerable effort has been made to improve predictions of runoff and erosion on rangelands through projects such as the Revised Universal Soil Loss Equation (RUSLE) (Renard *et al.,* 1991) and Water Erosion Prediction Project (WEPP) (Laflen *et al.,* 1991). However, estimation of model parameters for semiarid rangelands is often hampered by the spatial variability of erosion and hydrological processes.

The need to incorporate spatial variation of vegetation into hydrologic and erosion prediction models has been recognized (Blackburn *et al.,* 1992; Wilcox *et al.,* 1992). Spaeth *et al.* (1996) used multivariate statistical methods to stratify rangeland plant communities across landscapes that included nine western states. Their method increased the coefficient of determination (r^2) for regression equations developed to predict effective terminal infiltration rates.

The overall objective for this project was to determine if spatial variability of runoff and sediment yield for a long term study area could be stratified based on

[1]Paper No. 98181 of the *Journal of the American Water Resources Association.* **Discussions are open until February 1, 2002.**
[2]Respectively, Research Associate, Rangeland Ecosystem Science Department, Colorado State University, 1835 Parkview Blvd., Colorado Springs, Colorado 80906; Professor, Rangeland Ecosystem Science Department, Colorado State University, Fort Collins, Colorado 80523; Professor, Civil Engineering Department, University of Wyoming, Laramie, Wyoming 82070; and Hydrologist, Agricultural Research Service, Northern Plains Area Office, 1201 Oakridge, Suite 150, Fort Collins, Colorado 80525 (E-Mail/Mergen: demcjc@att.net).

plant life form within a shrub/grass plant community and to determine runoff and sediment yields from the initial 15 minutes of rainfall. This specific study was conducted to determine if site variability could be stratified based on results of a discriminant model of plant life form variables. Improved predictive capabilities of regression models and decreased standard errors would indicate stratification based on plant life form variables would be a successful method to stratify variability in runoff and sediment yield. The discriminant model could be incorporated into most hydrology and erosion models for other sites because it is based on plant life forms. The discriminant model developed in this study also provides a method to estimate how a change in management might affect hydrological processes. For example, reduction of the proportion of one group at a site should result in an increased proportion of another group. The effect on runoff and sediment yield could then be estimated by simply adjusting the weighted proportions of parameters randomly sampled for these two groups.

This study was designed to use observational data and to stratify spatial variability in surface hydrology for a sagebrush-grass rangeland using cluster and discriminant analysis. The hypothesis tested was that multiple regression equations developed for runoff and sediment yield for a single large data set should be equally robust compared with multiple regression equations that represent specific vegetation cluster groups resulting from cluster analysis. Specific objectives were to: (1) develop a discriminant model that could be applied at the beginning of most hydrological and erosion models that would stratify spatial variability of plant life form; (2) identify plant variables that best discriminated among the cluster groups; and (3) compare predictive capabilities of multiple regression equations and standard errors for each cluster group with predictive equations and standard errors for the regression models developed from the original data set.

The broad application of this study would be to use the discriminant model to stratify spatial variability in runoff and sediment yield for any hydrological model. Because the discriminant model is based on plant life forms, it could be used at any shrubland-grass site where canopy cover data were collected. Multiple regression equations are site and rainfall event specific and are only developed to test the hypothesis that this method can be used to stratify spatial variability based on plant life forms.

Study Area

The study was conducted during the 1993 and 1994 summer months (June through August) on the Sheep Creek allotment in the Roosevelt National Forest of Colorado (Figure 1). The study area was located 80 km northwest of Fort Collins, Colorado at 2,600 m elevation. Uplands in the Sheep Creek grazing allotment consist primarily of sagebrush-grass rangeland; the dominant shrub species is big sagebrush (*Artemisia tridentata* Nutt.), and dominant grasses are Idaho fescue (*Festuca idahoensis* Elmer), mountain muhly (*Muhlenbergia montana* (Nutt.) A.S. Hitchc.), Parry oatgrass (*Danthonia parryi* Scribn.), and needle grasses (*Stipa* spp.). Annual plant production was estimated at 500 kg/ha in unfavorable years to 1,000 kg/ha during favorable years (USDA, 1980).

The climate at Sheep Creek is temperate montane. Summer rainfall events recorded at the study site in 1993 and 1994 included two thunderstorms that produced more than 25 mm of precipitation. Although thunderstorms that produce more than 25 mm of rainfall are infrequent events at Sheep Creek, this type of storm produces runoff and sediment that enters the creek.

The shallow and well drained upland soils were in the Ratake soil series, classified as loamy-skeletal, mixed, shallow aridic, haploboroll. Soils near the riparian area were in the Naz 70 soil series, and classified as coarse loamy pachic cryoboroll (Pearce, 1995).

METHODS

One hundred plots with a range of cover from 0 to 89 percent for big sagebrush were selected. Fourteen plots were adjacent to the riparian zone at the uplands-riparian boundary and 86 were in the uplands.

Canopy cover, canopy height, surface roughness, and soil surface cover were measured for each plot using a 100-point pin table (Linse, 1992). The pin table was 0.6 m wide by 2 m long and fit within each plot of the same size. Each pin hit was used to determine canopy and soil surface cover, and to measure plant height and surface roughness. Each pin was lowered toward the ground surface one at a time. When the pin contacted plant canopy (first hit), the plant species and form was recorded as was plant height. The same pin was then lowered to the soil surface (second hit), and the second hit was recorded as soil surface cover or bare soil and an elevation of the soil surface was recorded. Canopy cover by species and soil surface cover were recorded as frequency of hits by the 100 points per plot. Surface roughness was calculated as a standard deviation of pin heights (Kuipers, 1957). Cover by plant life form and soil surface characteristics were categorized into six classes:

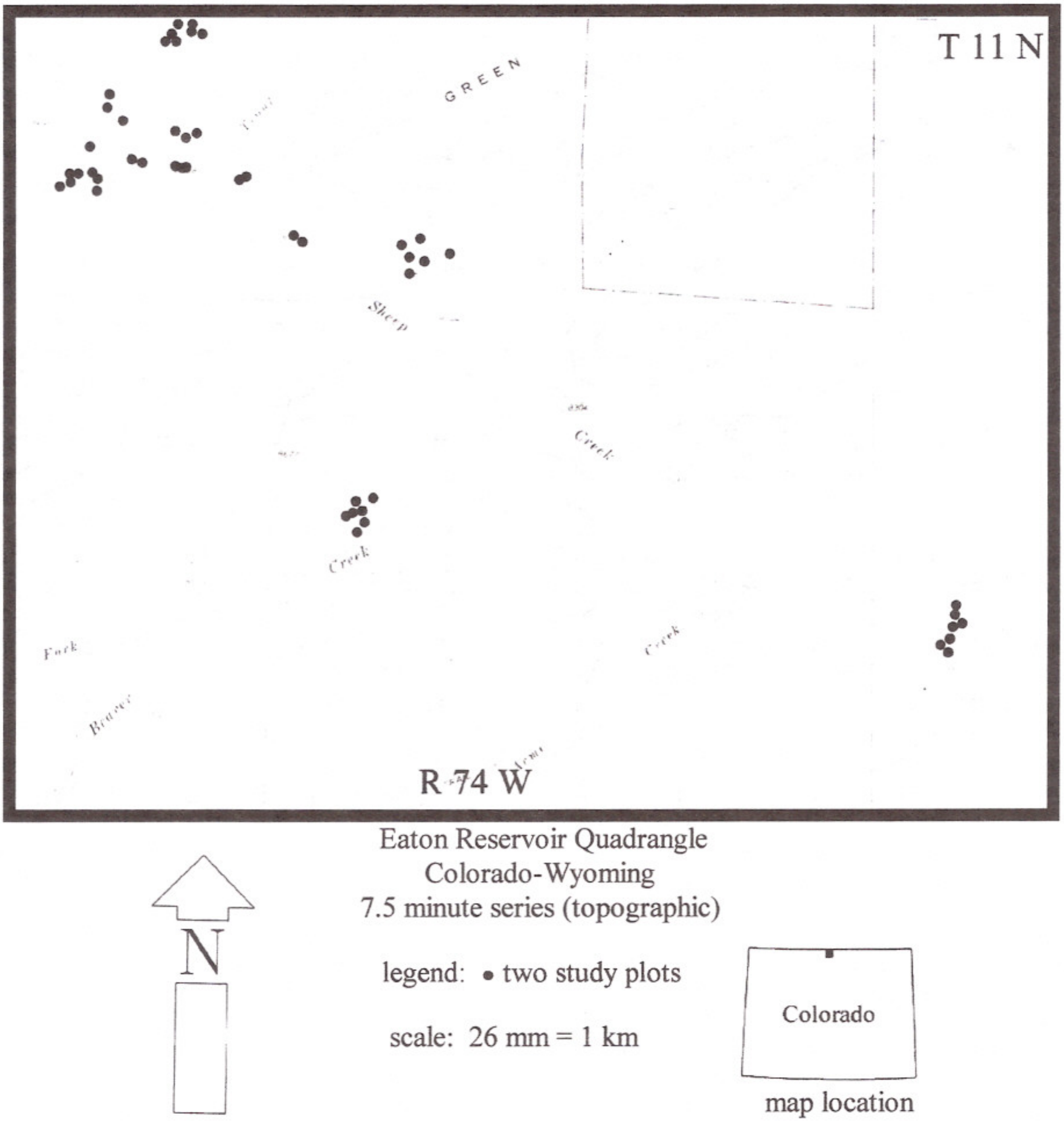

Figure 1. Location of Rainfall Simulation Plots at Sheep Creek, Colorado, Study Area. Each dot represents the approximate location of two study plots.

grasses, forbs, shrubs, cryptogams, bare ground, and litter.

Metal strips (15 cm high) were driven 3 cm into the soil on the top and both sides of each plot to delineate the 0.6 x 2 m plot boundary. A metal flume was installed at the lower end of each plot to funnel runoff into a plastic pipe and then into a collection bucket (Linse, 1992). Plot elevation and dimensions were collected with standard survey methods (Zeiss, 1986), and slope was calculated by rise/run.

A rotating boom rainfall simulator (Linse, 1992; Benkobi *et al.,* 1993, 1994; Pearce *et al.,* 1997) was used to simulate rainfall at an intensity of 100 mm/hr for 15 minutes at antecedent soil moisture conditions. This storm intensity was selected to insure runoff and sediment would be produced on most plots, hold the rainfall variability as constant as possible, and simulate a high intensity short duration thunderstorm that typically occurs on rangelands. A storm of this intensity has a return period of two to four years in southeast Wyoming (Huffsmith, 1988; Linse, 1992); there is a 25 percent chance that a storm of equal or greater rainfall will occur within five to six years 20 km south of the study site (estimated with 41 years of daily weather data for the months of May through September).

Two plots, separated by a distance of one meter, were utilized, and rainfall was applied simultaneously to both plots. Each plot was treated independently because plots were selected based on canopy cover of sagebrush. Generally the adjacent plot, which had to be one meter to the right or left of the first plot that was selected, seldom had similar plant cover or soil characteristics. Therefore, each plot was within a wide range of canopy cover variability, and often times, adjacent plots would be placed in different cluster groups.

Grab samples of initial runoff (when runoff first came off plots) were collected for 30 seconds and again at times of 5, 10, and 15 minutes after the rainfall began. The remaining runoff was collected in a bucket; this volume was added to the grab samples to calculate total runoff volume for each plot. Infiltration was calculated by subtracting the runoff amount from the applied rainfall (Thurow *et al.,* 1988); therefore, subsurface runoff through macropores, evapotranspiration, and water storage in the topography, litter, and vegetation were all combined as infiltration.

All grab samples and a final sample from the runoff in the bucket were filtered through preweighed Whattman glass microfiber filters to determine the mass of sediment (Pearce, 1995). Total sediment yield for each plot was calculated by adding the weight of sediment in the bucket and all grab samples.

Three 68.7 cm^3 soil samples were collected at a depth of three to six centimeters near each plot immediately prior to the start of the rainfall simulations. Antecedent soil moisture was determined by the gravimetric method (Gardner, 1986), bulk density by the core method (Blake and Hartge, 1986), and soil texture was determined by the hydrometer method (Bouyoucos, 1962; Allen, 1990) from the three soil samples adjacent to each plot. Additional soil samples were ashed in a muffle furnace at 550°C for five hours to determine soil organic matter (Storer, 1984).

Data Analysis

The two dependent variables were log-transformed because plots of residuals indicated nonconstant variance, residuals were not normally distributed, and plots of independent and dependent variables selected in regression equations prior to log transformations indicated some relationships other than linear existed (Ott, 1988). Stepwise multiple linear regressions were performed on the single data set with the log-transformed dependent variable. Diagnostic checks were made for outlying and influential cases, independent (x) observations, dependent (y) observations with studentized residuals, hat matrix leverage values, and covariance ratio; influential observations were identified with fitted values (Dffits), and the influence of each observation with each regression coefficient (Dfbetas) was checked. Coefficients of determination (r^2), significance of t-test for individual variables, and standard errors were checked for each equation and each variable. Collinearity was checked with variance inflation factors, condition numbers, variance-decomposition proportions, and the magnitude of regression coefficients and their signs with respect to realistic expectations (Belsley *et al.,* 1980; SAS, 1988; Belsley, 1991).

Cluster analysis is an empirical method that can be used quantitatively to group individuals, species, or objects into unknown groups. This multivariate procedure separates a single data set into groups that are not necessarily known before the analysis (Afifi and Clark, 1990). Cluster analysis was performed based on the percentage canopy cover for the six cover classes (grasses, forbs, shrubs, cryptogams, bare ground, and litter). Data were analyzed using a nonhierarchical cluster analysis procedure (ISODATA) to form groups based on group similarities within the group and among group differences (Ball and Hall, 1967; Hall and Khanna, 1977).

Stepwise multiple linear regression analyses (Ott, 1988) were performed to determine runoff and sediment yield relationships for the single data set and each vegetation cluster group. This process was used to select a set of independent variables based on cover and soil characteristics to be included in regression models as prediction variables (SAS, 1988).

Each of the cluster groups was considered a population of interest (Wester, 1992), and a general linear model (GLM) one-way analysis of variance (ANOVA) was performed to test for differences among cluster groups (Ott, 1988). Least significant difference tests (L.S.D.) were used to separate significantly different ($p \leq 0.05$ and 0.10) means (SAS, 1988). Coefficients of determination (r^2) and standard errors were used from the multiple regression analyses to display the amount of variation accounted for by each regression equation for comparison purposes. The amount of explained variation for the single data set was compared with results from multiple regression analyses of each cluster group. Differences (percent) between cluster groups and the single data set were also compared to determine if this technique improved predictions of runoff and total sediment yield.

A stepwise discriminant procedure was used to calculate standardized canonical discriminant function coefficients that were used to analyze differences among canopy cover variables simultaneously for cluster groups (Klecka, 1987). Therefore, information about numerous independent variables was contained within a single index (SPSS, 1986). Fisher classification coefficients (Klecka, 1987) were calculated to provide a method to assign (classify) any unclassified case (plot) into groups that they most closely resembled. The discriminant analysis also identified vegetation cover variables that contributed to the differences among the groups determined by the cluster analysis (Afifi and Clark, 1990). With this technique, the spatial variability that results from vegetation life forms can be stratified before building models to predict runoff, infiltration, and erosion for a plant community at most sites.

RESULTS AND DISCUSSION

Results of the diagnostics tests (SAS, 1988) for the 100 plots included in regression analysis indicated four cases (plots 33, 34, 36, and 38) greatly affected results of regression equations for the complete data set. Plots 33, 34, and 38 had studentized residuals much greater than |2|, the cutoff value. The criteria used for comparisons were based on sample size and number of regression variables. All four cases had a hat matrix value greater than the 0.08 criterion and Dffits values that were three to four times larger than the calculated criterion of 0.4. The covariance ratio was greater than expected for plots 36 and 38; all four cases influenced the intercept, and at least two of the regressors were identified as important, based on Dfbetas (Belsley *et al.,* 1980; Belsley, 1991). The four cases represented plots with little or no runoff during the rainfall simulation. This greatly influenced the regression analyses for the entire data set. Therefore, data for plots 33, 34, 36, and 38 were removed, and regression analysis was rerun and is discussed below. After stratifying the plots into the vegetation groups, these four plots were no longer identified as influential for the regression analysis and were therefore included in the analyses of the cluster groups.

Results of collinearity checks indicated there were no major problems with collinearity among any of the regression variables. The regression equations developed in our study were only used to test our hypothesis that the stratified data set based on plant forms would result in more robust regression equations when compared to the complete data set. Therefore, even with slight collinearity among variables like total cover and litter, which were expected with our use of observational variables, this did not have any harmful consequence to our hypothesis tested in this study.

Single Data Set and Runoff

The time that runoff began was selected first in the stepwise linear regression analyses for the entire data set and accounted for 48 percent of the variation in runoff (Table 1). The earlier that runoff began usually resulted in greater total runoff during a 15 minute rainfall simulation. The time that runoff began integrated several unmeasured variables (micro-channel geometry, channel density, and tortuosity) as well as measured variables (slope, soil moisture, organic matter, bulk density, texture, cover, and roughness) into one independent variable. These variables reflect conditions at or near the soil surface and have the greatest influence on runoff and infiltration early in a rainstorm. Time runoff began was correlated with total runoff ($r = -0.70$, $p < 0.01$), total ground cover ($r = -0.19$, $p = 0.06$), and silt ($r = 0.24$, $p = 0.02$). Therefore, the unmeasured variables may be more important than measured variables to determine runoff at antecedent soil moisture conditions.

Soil texture and soil organic matter (SOM) were selected as the next most important variables, each contributing 2 to 3 percent of the total explained variation for runoff. Infiltration is affected by soil organic matter because an increase in soil organic matter in turn binds soil aggregates (Blackburn, 1975; Takar *et al.,* 1990), and greater soil organic matter and aggregates have been shown to be positively correlated with infiltration rate (McCalla *et al.,* 1984a). Silt had a positive coefficient which indicated that the greater the silt content, the greater the runoff. Clay had a negative coefficient indicating less runoff. This observation was counterintuitive, but percent clay was very low for soils at the study site (3 to 7 percent) and a greater clay content (a difference between 3 to 7 percent) indicated less runoff. Plots with greater clay content soils were found on grass and shrub covered plots which often had greater runoff.

The coefficient of determination (r^2) increased from 0.24 to 0.56 when data for four plots that had little or no runoff during rainfall simulations were removed from the analyses. These four plots were located near each other and had a very shallow soil with a gravel texture, which probably allowed infiltration to occur at a high rate (Tromble *et al.,* 1974; Seyfried, 1991). These four plots were not represented by extreme values of vegetation cover or soil characteristics, but they did represent plots that could be expected to produce little runoff at existing antecedent soil moisture.

TABLE 1. Stepwise Multiple Linear Regression Equations for the Single Data Set and for Cluster Groups. Variables are listed in the order they were selected in the stepwise procedure. Groups were formed using cluster analysis and are identified by the dominant vegetation life form.

Predictive Equation	N	Multiple r^2	Standard Error	Runoff or Sediment Yield
Single Data Set				
y = 38.67 - 2.66*Time + 0.26*Silt - 1.19*SOM - 0.89 Clay	96	0.56	4.13	Runoff
y = 4.45 - 0.18*Time - 0.02*Totveg - 0.03*Crypt - 0.02*Litter	96	0.27	0.83	Sediment Yield
Grass Group				
y = 44.48 - 3.01*Time - 0.51*SM - 0.12*SR	40	0.65	4.09	Runoff
y = 1.69 + 0.04*Baresoil2 + 0.01*Litter2 - 0.03*Totveg	40	0.43	0.62	Sediment Yield
Grass/shrub Group				
y = 36.81 - 1.94*Time - 0.4*Forb + 0.26*Litter	23	0.67	3.45	Runoff
y = 4.30 - 0.41*Time - 0.08*Forb	23	0.40	1.06	Sediment Yield
Shrub Group				
y = 25.55 - 1.67*Time + 0.54*Silt - 2.08*SOM	19	0.68	3.02	Runoff
y = -0.35 - 0.14*SM + 0.12*Silt - 0.45*SOM + 0.06*Forb	19	0.91	0.27	Sediment Yield
Forb/Grass Group				
y = 100.17 - 94.63*BD + 1.56*Baresoil2 + 0.47*Grass	18	0.73	6.70	Runoff
y = -1.33 + 0.04*SR + 0.03*Totveg	18	0.37	0.79	Sediment Yield

Notes:
1. Dependent variable (y) is total runoff after being transformed (log (ml +1)) or total sediment after being transformed (log (g + 1)).
2. Independent variables: Time = the time runoff began (min); Silt = silt present in the soil (%); SOM = soil organic matter content in the soil (%); Clay = clay present in the soil (%);Totveg = total canopy cover (%); Crypt = cryptogam cover (%); Litter = litter cover (%); SM = antecedent soil moisture (%); SR = surface roughness (std. dev.); Baresoil2 = bare soil (%) recorded on second hit of pins; Litter2 = litter cover (%) recorded on second hit of pins; Forb = forbs cover (%); BD = soil bulk density (g cm^{-3}); Grass = grass cover (%).

Single Data Set and Sediment Yield

Variables that represented total vegetation, cryptogam, litter cover (first pin hit), and the time runoff began were selected as important variables for sediment yield (Table 1). The time runoff began was the first variable selected and explained the most variability (13 percent) in sediment yield. This variable also had a negative coefficient, which indicated that the later runoff began, the less sediment that was eroded from plots. Time that runoff began was only correlated with sediment yield ($r = -0.36$, $p < 0.01$). The lack of any significant correlation of the time that runoff began with measured variables indicates that unmeasured variables may be more important to explain sediment yield than are the measured variables.

Total vegetation cover, cryptogam cover, and litter, which protected the soil from raindrop impact, all reduced sediment yield and explained about 15 percent of the variation in sediment yield during a 15 minute rainfall simulation. Conversely, the more bare soil susceptible to raindrop impact, the greater the sediment yield.

The four plots where little or no runoff occurred, which significantly affected the results of regression analysis when runoff was the dependent variable, were determined as being only moderately important in the analysis of sediment yield. The coefficient of determination (r^2) increased from 0.22 to 0.27 when these four plots were excluded from the regression analyses.

Data Separated Into Four Vegetation (cluster) Groups

The 100 plots were separated into four cluster groups based on differences in vegetation composition and life form. These groups stratified the effects that vegetation type and canopy cover had on the soil and soil surface, which in turn affected runoff, infiltration, and sediment yield from the plots. The amount and type of canopy cover, litter, and bare ground within each of the four vegetation groups was similar with a minimum variance within each group (del Morel, 1975). Cluster analysis maximized differences among cluster groups (Ratliff and Pieper, 1981). Each cluster group represented a specific range of canopy cover representative of variability at the study site. The four groups were based on the dominant vegetation cover and life form and were classified as: (1) grass, (2) grass/shrub, (3) shrub, and (4) forb/grass. Group means and significant differences found for dependent and independent variables were calculated (Table 2).

Stepwise Multiple Regression Analyses of Groups

Results of the stepwise multiple linear regression analyses for each cluster group with total runoff and total sediment yield as dependent variables are shown in Table 1. Variables are listed in the order they were selected in the regression analyses and all variables were significant at $p \leq 0.10$. Classification of plots, based on vegetation form and canopy cover, improved r^2 values and thus predictive capability of the regression equations in all cases. The standard error was less for the grass, grass/shrub, and shrub groups compared to the single data set when runoff was the dependent variable. The grass/shrub group was the only group with a greater standard error than the single group when sediment yield was compared. Therefore, stratification of the single data set resulted in smaller standard errors in three of four regression equations for both runoff and sediment yield and resulted in more robust predictive equations for all four groups.

Runoff

The coefficients of determination (r^2) for the runoff predictive equations for the four groups ranged from 0.65 to 0.73 (Table 1). In most analyses, the time when runoff began was the most important variable and explained as much as 54 percent of the variability in total runoff. Cover variables and soil variables such as surface roughness, soil moisture content, silt content, and soil bulk density were also important variables and accounted for 4 to 24 percent of the variation in total runoff.

Total runoff was significantly correlated with the time runoff began in all groups except the forb/grass group. The time that runoff began was only correlated with silt ($r = 0.28$, $p = 0.08$) in the grass group, antecedent soil moisture ($r = 0.39$, $p = 0.10$) in the shrub group, and grass cover ($r = 0.48$, $p = 0.04$) in the forb/grass group. These correlations indicate the time runoff began may be influenced by these variables, but the unmeasured variables may be of greater importance to explain runoff.

Results from a GLM one-way ANOVA indicated that total runoff was significantly lower from plots classified into the forb/grass group as compared with the other three groups ($p \leq 0.05$). When the four plots that had little or no runoff were removed for the one-way ANOVA, no differences in runoff were found among the four vegetation groups. However, these four plots were not identified from the diagnostic checks as influential data when analyzed within the forb/grass group for regression analysis. Significant differences were found among the four vegetation groups in soil bulk density, organic matter, surface roughness, and plant height as well as many canopy cover variables (Table 2).

The grass cluster group had more initial runoff volume (0.28 $l\ min^{-1}m^{-2}$) than the other three groups, and runoff began earlier (4.6 minutes) compared with the grass/shrub and forb/grass groups, but because infiltration was high, less total runoff volume (4.4 L) was produced. Runoff was underpredicted by 15 percent in the initial sample and overpredicted by 11 percent during the final sample when compared with the single data set.

Runoff from the grass/shrub group and the forb/grass group did not reach equilibrium because surface microtopography on plots created a soil surface with a very nonuniform infiltration rate. The nonuniformity of infiltration was also caused by differences in the amounts and forms of canopy cover. For example, shrubs and litter created obstacles for runoff and increased the intake opportunity time for infiltration. The amount of bare soil and greater bulk densities decreased infiltration, but grasses and forbs increased infiltration. This array of variables influenced antecedent soil moisture, flow paths, and infiltration rates. Interactions among these variables with the gradual increase in runoff depth for the short duration rainfall simulations prevented runoff equilibrium.

The shrub group had the greatest canopy cover and height, litter cover, and surface roughness (Table 2). This resulted in increased resistance to overland flow and caused a five-minute delay in runoff equilibrium when compared with the grass group. Runoff from plots in the shrub group had a greater effective distance to travel than for the other three groups, however total runoff volume was greater (5.8 L, although

TABLE 2. Means and Standard Errors of the Mean for All Dependent and Independent Variables and the Results of One-Way Analysis of Variance Among Cluster Groups.

Variable	Grass Cluster Group Mean n = 40	SE	Grass/Shrub Cluster Group Mean n = 23	SE	Shrub Cluster Group Mean n = 19	SE	Forb/Grass Cluster Group Mean n = 18	SE	p
Sediment Yield (g)	1.37 a[1]	0.23	2.27 a	0.65	1.92 a	0.32	1.52 a	0.42	< 0.40
Runoff (ml)	4,386 a	651	4,004 a	851	5,753 a	1,164	3,047 b	661	< 0.01[2]
Bulk Density (g cm^{-3})	1.09 bc	0.02	1.16 ab	0.02	1.07 c	0.03	1.19 a	0.02	< 0.01
	1.09 b	0.02	1.16 a	0.02	1.07 b	0.03	1.19 a	0.02	< 0.10
Surface Roughness (std. dev. mm)[3,5]	12 b	2	19 b	2	38 a	4	20 b	4	< 0.01
	12 c	2	19 b	2	38 a	4	20 b	4	< 0.10
Soil Organic Matter (% by weight)	3.5 a	0.2	2.6 b	0.1	2.8 b	0.2	2.8 b	0.1	< 0.01
Plant Height (mm)[4]	167 c	16	238 b	16	368 a	27	163 c	21	< 0.01
Time for Runoff to Begin (min)	4.6 a	0.3	5.2 a	0.3	4.6 a	0.4	4.9 a	0.8	< 0.70
Grass Cover (%)[4]	56 a	2	35 b	2	16 d	2	27 c	2	< 0.01
Forb Cover (%)[4]	27 a	3	11 b	1	5 b	1	21 a	3	< 0.01
Shrub Cover (%)[4]	2 c	1	27 b	3	68 a	3	4 c	1	< 0.01
Cryptogams Cover (%)[4]	4 a	1	3 a	1	1 a	< 1	6 a	2	< 0.15
Litter Cover (%)[4]	8 b	1	8 b	2	7 b	1	36 a	3	< 0.01
Bare Ground (%)[4]	3 b	1	16 a	4	3 b	1	7 b	1	< 0.01
Bare Ground (%)[5]	20 ab	3	27 a	5	6 c	2	12 b	1	< 0.01
Litter (%)[5]	58 b	4	60 b	5	86 a	2	63 b	3	< 0.01
Total Aerial Cover (%) (grass+shrub+forb+cryptogam)[4]	89 a	1	76 b	4	91 a	1	57 c	4	< 0.01
Total Ground Cover (%) (cryptogam+litter+live plants)[5]	80 bc	3	73 c	5	94 a	2	88 ab	1	< 0.05
	80 b	3	73 c	5	94 a	2	88 a	1	< 0.10
Sand Content (%)	65 a	1	68 a	2	67 a	2	65 a	2	< 0.50
Silt Content (%)	30 a	1	27 a	1	29 a	2	31 a	2	< 0.40
Clay Content (%)	5 a	< 1	5 a	< 1	4 a	< 1	5 a	< 1	< 0.40
Antecedent Soil Moisture (%)	6.2 a	0.6	5.1 a	0.7	6.3 a	0.9	4.8 a	0.8	< 0.40
Slope (%)	15 a	1	16 a	< 0.5	15 a	1	14 a	1	< 0.70

[1]Means in a row followed by a similar letter are not significantly different at the listed p-values in the last column.
[2]Runoff was not found to be significantly different when influential data (plots 33, 34, 36, and 38) were removed from the analysis.
[3]Surface roughness was calculated as a standard deviation of the mean and the SE is the SE of the mean.
[4]Data collected on the first hit of each pin.
[5]Data collected on the second hit of each pin.

not significantly). Runoff at 15 minutes was 14 percent greater than the average sample calculated from the single data set. The final runoff rates at the end of rainfall simulations were slightly different among all groups, being the lowest in the grass (0.60 L/min/m^2), followed by the, grass/shrub (0.67 L/min/m^2), forb/grass (0.74 L/min/m^2) and greatest (0.78 L/min/m^2) in the shrub group.

The lack of differences observed among groups in runoff was greatly influenced by the variability in runoff within the grass/shrub and forb/grass groups, which resulted in runoff equilibrium not being reached within the 15-minute rainfall simulations. These two groups had the greatest variability in total aerial and ground canopy cover, litter, cryptogams, and bare ground. This variability in vegetation affected soil surface properties like surface roughness, tortuosity, surface storage space for runoff, channel geometry and density, in addition to antecedent soil moisture conditions. Vegetation also influenced the amount of rainfall intercepted and evapotranspiration by the different amounts and types of vegetation, and this would influence the antecedent soil moisture. The variability of soil and vegetation characteristics within groups influenced the amount of runoff during the first 15 minutes of the rainfall simulation so that significant differences could not be detected with ANOVA. All four groups had different runoff rates at the end of 15 minutes. Significant differences in runoff among groups may have been observed if runoff equilibrium had been reached for all groups or if the rainfall simulation period had been extended beyond 15 minutes. Separation of the data into the vegetation groups did stratify the variability of many variables (Table 2) into groups that are similar within a group and different among groups. Runoff predictions for similar rainfall intensity and duration may be adequate without group separation. However, clustering plots into groups with similar plant and soil characteristics which do influence runoff, may improve the prediction of runoff, especially for longer rainfall periods or with soils that have greater antecedent soil moisture.

Sediment Yield

Predictions of sediment yield were better for each of the four vegetation groups as compared to the single data set. The multiple r^2 for the predictive equations for each of the four groups ranged from 0.37 to 0.91 compared to the equation (r^2 = 0.27) for the single data set (Table 1).

The most important single variable among the four different groups for most multiple regression analyses was the type of vegetation cover. Total vegetation cover was selected in two groups, however litter and forb cover and bare soil were also selected as important variables (Table 1). Time for runoff to began was only selected as important to predict sediment yield in the grass/shrub group, but was not correlated with any other independent variables. Soil variables such as soil moisture content, soil organic matter, and silt content were found to be somewhat important. Surface cover variables explained 5 to 28 percent of the variability in sediment yield, whereas soil characteristics explained from 18 to 34 percent of the variation in sediment yield.

Results from a GLM one-way ANOVA indicated that total sediment yield was similar among groups. This result was partially due to the fact the grass/shrub and forb/grass groups did not reach runoff equilibrium within the 15-minute rainfall simulation. However, comparing the total sediment yield within each group to the average for the original data set indicated total sediment yield would be overpredicted in the grass group by 20 percent and by 11 percent in the forb/grass group. Total sediment yields would be underpredicted by 25 percent in the grass/shrub group and 11 percent in the shrub group. Therefore, the small scale spatial variability found on different plots was effectively stratified based on vegetation life form and bare ground, and stratification offers a way to improve predictive capabilities of hydrology and erosion equations within a plant community. This stratification is based on vegetation life forms which have been identified as being a great influence to the spatial variability of hydrological processes.

Vegetation directly influences the quantity of sediment that is lost from a plot. The vegetation intercepts raindrops to reduce the impact they have on the soil surface, thus less sediment dislodged to be eroded. Plant height of the various forms can influence the amount of airborne sediment that has accumulated on or under vegetation. The indirect influence would be how vegetation affects runoff as discussed above. Less runoff would probably transport less sediment off plots, and greater surface storage, tortuosity, small dams created by litter could influence the amount of sediment in runoff that is redeposited on a plot. The lack of differences observed in sediment yield among groups was influenced by vegetation canopies and the amount of runoff per group. At the end of the 15-minute rainfall, sediment was eroding off plots at a rate of 0.20 g/min/m^2 in the grass group, 0.52 g/min/m^2 for the grass/shrub, 0.22 g/min/m^2 for the shrub, and 0.27 g/min/m^2 for the forb/grass group. Some of these differences were influenced by the grass and shrub groups approaching runoff equilibrium, and the other two groups were not near runoff equilibrium after 15-minute rainfall simulations. Different hydrological processes were occurring among groups near the end of the rainfall simulation. The grass and shrub groups were functioning under hydrological processes associated with runoff equilibrium and processes for the grass/shrub and forb/grass groups like infiltration, ponding, and flow paths being connected were approaching a maximum rate or completion. Therefore, separation of data into groups may improve predicted sediment yields during a 15-minute rainfall period from our study site, and significant differences in sediment yield may be found if the rainfall duration was extended or if antecedent soil moisture was greater; this would allow all plots to reach runoff equilibrium.

Discriminant Analysis

The standardized canonical discriminant function coefficients (Klecka, 1987) indicated the relative importance each variable had within each function (Table 3). It was possible to determine the importance that each variable contributed to the functions that separated the groups by examining the magnitude of the standardized canonical discriminant function coefficients. Four variables (cover of grasses, shrubs, litter, and bare ground) were required to estimate the three canonical discriminant functions. Shrub cover, which accounted for 81 percent of the total variation, was the most important variable in the first function for group separation. Shrubs have a large effect on the microenvironment and have been shown to influence rangeland hydrology through the formation of coppice dunes (Blackburn, 1975; Pierson *et al.,* 1994). Litter and grass cover were important for Function 2, and both have been found to be very important in affecting infiltration and runoff (Meeuwig, 1970; McCalla *et al.,* 1984b). Bare ground was important in Function 3 and often has been used to explain sediment yield (McCalla *et al.,* 1984a; Warren *et al.,* 1986). Therefore, of the variables entered in the discriminant analysis, the cover of shrubs, grasses, litter, and bare ground were most important for discriminating among the four vegetation cluster groups. Cover of forbs and cryptogams, and variables recorded on the second pin hit (bare ground, litter, and total ground cover) were not required or selected as being important in the discriminant analysis.

The magnitude of the Fisher's classification function coefficients indicated the importance that each variable had among groups (Table 4). Grasses and bare ground contributed most to classification of plots in the grass group. Shrubs and bare ground contributed most to the grass/shrub group, and shrubs were most important for classification into the shrub group. Litter contributed the most for the forb/grass group while, bare ground, grasses, and shrubs all contributed similarly in importance. The contribution of each variable among the four groups reflected the vegetation composition of each group.

The Fisher's classification function coefficients can be used for classification of new plots and could be included into hydrology and erosion models (Table 4). This model could be used to separate the spatial variability runoff and sediment yield caused by vegetation into one of four groups prior to runoff and sediment yield estimates at any shrubland-grass site. A posterior classification correctly placed 97 percent of the plots in correct groups (SPSS, 1986) and a jackknife procedure classified 95 percent of the plots correctly (SAS, 1988).

TABLE 3. The Standardardized Canonical Discriminant Function Coefficients With Variables Listed in Order of Entrance Into the Discriminant Analysis. The size of the coefficient (absolute value) is an indication of the importance the variable has within a function.

Variables	Function 1	Function 2	Function 3
Bare Ground	0.57	-0.14	1.15
Grasses	0.04	-0.56	0.31
Litter	-0.03	0.82	0.16
Shrubs	1.15	-0.05	0.12
Percent of the Variation Explained by Each Function	81	17	2

The discriminant model could easily be incorporated into rangeland hydrology and watershed models. The Fisher's classification function coefficients could be programmed at the beginning of most hydrological models if canopy cover data for a model were used as input data. The percent canopy cover for bare ground, grasses, litter, and shrubs would be used in the discriminant model to determine to which group the data belongs. Once the group is determined, the model would continue to run through the remainder of the computations and result in predicted runoff and

TABLE 4. Fisher's Classification Function Coefficients for Group Classification of Rangeland Plots Near Sheep Creek, Colorado. The magnitude of each coefficient for each group indicates the importance of this variable for classification among groups.

Variables	Grass Group	Grass/Shrub Group	Shrub Group	ForbGrass Group
Bare Ground (%)	0.45	0.77	0.87	0.43
Grasses (%)	0.57	0.57	0.52	0.37
Litter (%)	0.25	0.34	0.36	0.71
Shrubs (%)	0.38	0.92	1.54	0.38
(Constant)	-19.37	-31.31	-60.70	-21.29

Note: To classify a new plot into a group, multiply the percent canopy cover for each of the 4 variables (bare ground, grasses, litter and shrubs) by the coefficients in each row. (Example: bare ground (%)*0.45, bare ground (%)*0.77 etc., across the entire row). After multiplying (%) of each variable by each of the coefficients in that respective row, add each product by column (for each group) and add the constant. (Example: (% bare ground*0.45) + (% grasses*0.57) + (% litter*0.25) + (% shrubs*0.38) + (-19.37). Compare the resulting four numbers (1 number under each group), the largest of the four numbers indicates the classification of the new plot. This model could be tested at any shrubland-grass site for classification and then be incorporated into hydrology and erosion models.

sediment yields per group. This method could be tested with current plot data from other studies if data were collected during the initial 15 minutes of simulated rainfall.

CONCLUSION

Cluster analysis resulted in four vegetation groups called the grass, grass/shrub, shrub, and forb/grass groups. A discriminant model developed indicated the canopy cover of bare ground, grasses, litter, and shrubs can classify plots from the single data set (100 plots) into one of the four vegetation cluster groups. Classifying plots of the single data set into these four vegetation groups effectively stratified spatial variability based on plant life forms. This model can be used to classify data from other sites into one of four groups and could be used in most rangeland hydrological models to make this stratification before predicting infiltration, runoff, and sediment yield. Randomly sampling an area for site parameters and then classifying these samples using the discriminant model would automatically weight the proportion of samples classified into each of the four vegetation groups.

This method increased the amount of explained variation (measured as an increase in r^2) in total runoff and sediment yield for each of the four groups compared with the single data set. Standard errors of regression equations developed were also improved within three of the four groups for both runoff and sediment yield when compared with the single data set. This method could improve predictive capabilities for runoff, infiltration, and sediment yield for short duration rainfall events on other rangelands.

The time that runoff began was a variable we measured and used as an independent variable for regression analyses. The time that runoff began was the most important variable of the multiple regression equations developed in this study, which was a combination of measured (antecedent soil moisture, slope, organic matter, bulk density, texture, cover, and roughness) and unmeasured variables (micro-channel geometry, channel density, and tortuosity). The variable for the time that runoff began was slightly correlated with only a few measured variables; therefore, the importance of this variable is probably best explained with the unmeasured variables. This one variable, which explained much of the variability in runoff and sediment yield, needs to be studied further to determine the most important characteristics that are important for short duration rainstorms on rangelands. A better understanding of the unmeasured variables could help improve model predictions at antecedent conditions.

LITERATURE CITED

Afifi, A. A. and V. Clark, 1990. Computer-Aided Multivariate Analysis (Second Edition). Van Nostrand Reinhold Company, New York, New York.

Allen, T., 1990. Particle Size Measurement (Fourth Edition). Chapman and Hall, New York, New York.

Ball, G. H. and D. J. Hall, 1967. A Clustering Technique for Summarizing Multivariate Data. Behavioral Sci. 12:153-155.

Belsley, D.A., 1991. Conditioning Diagnostics Collinear and Weak Data in Regression. John Wiley and Sons, New York, New York.

Belsley, D. A., E. Kuh, and R. E. Welsch, 1980. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. John Wiley and Sons, New York, New York.

Benkobi, L., M. J. Trlica, and J. L. Smith, 1993. Soil Loss as Affected by Different Combinations of Surface Litter and Rock. J. Environ. Qual. 22:657-661.

Benkobi, L., M. J. Trlica, and J. L. Smith, 1994. Evaluation of a Refined Surface Cover Subfactor for Use in RUSLE. J. Range Manage. 47:74-78.

Blackburn, W. H., 1975. Factors Influencing Infiltration and Sediment Production of Semiarid Rangelands in Nevada. Water Resour. Res. 11:929-937.

Blackburn, W. H., F. B. Pierson, C. L. Hanson, T. L. Thurow, and A. L. Hanson, 1992. The Spatial and Temporal Influence of Vegetation on Surface Soil Factors in Semiarid Rangelands. Trans. ASAE 35:479-486.

Blackburn, W. H. and F. B. Pierson, Jr., 1994. Sources of Variation in Interrill Erosion on Rangelands. *In:* Variability of Rangeland Water Erosion Processes, W. H. Blackburn, F. B. Pierson, G. E. Schuman, and R. Zartman (Editors). SSSA Spec. Pub. 38. Madison, Wisconsin, pp. 1-9.

Blake, G. R. and K. H. Hartge, 1986. Bulk Density. *In:* Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods (Second Edition), A. Klute (Editor). Agron. Amer. Soc. Agron., Madison, Wisconsin, No. 9, pp. 363-375.

Bouyoucos, G. J., 1962. Hydrometer Method Improved for Making Particle Size Analysis of Soil. Agron. J. 54:464-465.

del Morel, R., 1975. Vegetation Clustering by Means of ISODATA: Revision by Multiple Discriminant Analysis. Vegetation 29:179-190.

Devaurs, M. A. and G. F. Gifford, 1984. Variability of Infiltration Within Large Runoff Plots on Rangelands. J. Range Manage. 37:523-528.

Gardner, W. H., 1986. Water Content. *In:* Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods (Second Edition), A. Klute (Editor). Agron. Amer. Soc. of Agron., Madison, Wisconsin, No. 9, pp. 493-544.

Hall, D. J. and D. Khanna, 1977. The ISODATA Method of Computation for the Relative Perceptions of Similarities and Differences in Complex and Real Data. *In:* Statistical Methods for Digital Computers, K. Enslein, A. Ralston, and H. W. Wilf (Editors). John Wiley and Sons, New York, New York, Vol. 3, pp. 340-373.

Huffsmith, R. L., 1988. Alternative Sediment Control in Surface Mine Reclamation. M.S. Thesis, Univ. of Wyoming., Laramie, Wyoming.

Klecka, W. R., 1987. Discriminant Analysis. Sage University Paper Series Quantitative Applications in the Social Sciences, Series No. 07-019. Sage Publ. Beverly Hills, California.

Kuipers, H., 1957. A Reliefmeter for Soil Cultivation Studies. Neth. J. Agric. Sci. 5:255-262.

Laflen, J. M., L. J. Lane, and G. R. Foster, 1991. WEPP, A New Generation of Erosion Prediction Technology. J. Soil Water Conserv. 46:34-38.

Linse, S. J., 1992. The Influence of Ground Cover on Upland Range Erosion. M.S. Thesis, Univ. of Wyoming., Laramie, Wyoming.

Meeuwig, R. O., 1970. Infiltration and Soil Erosion as Influenced by Vegetation and Soil in Northern Utah. J. Range Manage. 23:185-188.

McCalla II, G. R., W. H. Blackburn, and L. B. Merrill, 1984a. Effects of Livestock Grazing on Sediment Production, Edwards Plateau of Texas. J. Range Manage. 37:291-294.

McCalla II, G. R., W. H. Blackburn, and L. B. Merrill, 1984b. Effects of Livestock Grazing on Infiltration Rates, Edwards Plateau of Texas. J. Range Manage. 37:265-269.

Ott, L., 1988. An Introduction to Statistical Methods and Data Analysis (Third Edition). PWS-Kent Publ. Co., Boston, Massachusetts.

Pearce, R. A., 1995. Sediment Movement and Filtration Within Laboratory and Riparian Vegetation Buffer Strips. Ph.D. Diss., Colorado State Univ., Fort Collins, Colorado.

Pearce, R. A, M. J. Trlica, W. C. Leininger, J. L. Smith, and G. W. Frasier, 1997. Efficiency of Grass Buffer Strips and Vegetation Height on Sediment Filtration in Laboratory Rainfall Simulations. J. Environ. Qual. 26:139-144.

Pierson, F. B., W. H. Blackburn, S. S. Van Vactor, and J. C. Wood, 1994. Partitioning Small Scale Spatial Variability of Runoff and Erosion on Sagebrush Rangeland. Water Resour. Bull. 30:1081-1089.

Ratliff, D. and R. D. Pieper, 1981. Deciding Final Clusters: An Approach Using Intra- and Intercluster Distances. Vegetation 48:83-86.

Renard, K. G., G. R. Foster, G. A. Weesies, and J. P. Porter, 1991. RUSLE, Revised Universal Soil Loss Equation. J. Soil Water Conserv. 46:30-33.

SAS Institute Inc., 1988. SAS/STAT User's Guide, (Release 6.03 Edition). SAS institute, Cary, North Carolina.

Seyfried, M. S., 1991. Infiltration Patterns From Simulated Rainfall on a Semiarid Rangeland Soil. Soil Sci. Soc. Am. J. 55:1726-1734.

Simanton, J. R., M. A. Weltz, and H. D. Larson. 1991. Rangeland Experiments to Parameterize the Water Erosion Prediction Project Model: Vegetation Canopy Cover Effects. J. Range Manage. 44:276-282.

Spaeth, K. E., F. B. Pierson, M. A. Weltz, and J. B. Awang, 1996. Gradient Analysis of Infiltration and Environmental Variables as Related to Rangeland Vegetation. Trans. ASAE 39:67-77.

SPSS Inc., 1986. SPSSx User's Guide (Second Edition). Chicago, Illinois.

Storer, D. A., 1984. A Simple High Volume Ashing Procedure for Determining Soil Organic Matter. Comm. Soil Sci. Plant Anal. 15:759-772.

Takar, A. A., J. P. Dobrowolski, and T. L. Thurow, 1990. Influence of Grazing, Vegetation Life-Form, and Soil Type on Infiltration Rates and Interrill Erosion on a Somalion Rangeland. J. Range Manage. 43:486-490.

Thurow, T. L., W. H. Blackburn, and C. A. Taylor, Jr., 1988. Infiltration and Interrill Erosion Responses to Selected Livestock Grazing Strategies, Edwards Plateau, Texas. J. Range Manage. 41:296-302.

Tromble, J. M., K. G. Renard, and A. P. Thatcher, 1974. Infiltration for Three Rangeland Soil-Vegetation Complexes. J. Range Manage. 27:318-321.

USDA, Soil Conservation Service and Forest Service, 1980. Soil Survey of Larimer County Area, Colorado. U.S. Govt. Print. Off. 239-812/48, Washington, D.C.

Warren, S. D., W. H. Blackburn, and C. A. Taylor, Jr., 1986. Soil Hydrologic Response to Number of Pastures and Stocking Density Under Intensive Rotation Grazing. J. Range Manage. 39: 500-504.

Wester, D. B., 1992. Viewpoint: Replication, Randomization, and Statistics in Range Research. J. Range Manage. 45:285-290.

Wilcox, B. P., M. Sbaa, W. H. Blackburn, and J. H. Milligan, 1992. Runoff Prediction From Sagebrush Rangelands Using Water Erosion Prediction Project (WEPP) Technology. J. Range Manage. 45:470-474.

Zeiss, Carl, 1986. Zeiss Elta 3 Operating Instructions. Nominal Accuracy: Horizontal Angle Measurement, ± 2 Seconds; Distance Measurement, ± 3 mm with + 3 ppm. Zeiss. Oberkocher, West Germany.