Technical Report No. 83

SITE COMPARISONS OF ABOVEGROUND PLANT BIOMASS, 1970 SEASON

William E. Grant

Natural Resource Ecology Laboratory

Colorado State University

Fort Collins, Colorado

GRASSLAND BIOME

U. S. International Biological Program

April 1971

TABLE OF CONTENTS

ABSTRACT

Intersite comparisons of live aboveground plant biomass were made using data from the 1970 field season. The nine sites compared were the following: Bison, Bridger, Dickinson, Cottonwood, Pawnee, Hays, Osage, Pantex, and Jornada.

INTRODUCTION

This report presents a comparison of the floral composition (live above-ground biomass) at nine U.S. IBP Grassland Biome research sites. The sites represent various grassland habitats which differ in precipitation, elevation, length of growing season, and other important environmental parameters. Fig. 1 is a map showing the geographical location of the sites, and some pertinent site characteristics are listed in Table 1. Additional information concerning the individual sites can be obtained from the following Grassland Biome technical reports: Tech. Rep. No. 37 (Bison), Tech. Rep. No. 38 (Bridger), Tech. Rep. No. 39 (Cottonwood), Tech. Rep. No. 40 (Dickinson), Tech. Rep. No. 41 (Hays), Tech. Rep. No. 43 (Jornada), Tech. Rep. No. 44 (Osage), Tech. Rep. No. 1 (Pawnee), and Tech. Rep. No. 45 (Pantex).

The sites were compared using two different indices of similarity. The first index involved an information equation derived from Shannon and Weaver (1949) as described by Horn (1966). The second index was based on Jaccard's coefficient of community (after Greig-Smith 1964).

SHANNON AND WEAVER'S R_o EQUATION

The equation is written below:

$$R_o = \frac{\Sigma(x_i + y_i)\ \ln(x_i + y_i) - \Sigma\ x_i\ \ln\ x_i - \Sigma\ y_i\ \ln\ y_i}{(X + Y)\ \ln(X + Y) - X\ \ln\ X - Y\ \ln\ Y}$$

R_o is an index of overlap which, when the x_i and y_i values are expressed as proportions of the total X and total Y, varies from 0 when the samples are completely distinct (no species or groups in common) to 1 when the samples are identical with respect to proportional species or group composition (Horn 1966). When data are available in the form of frequencies, as is the case in this study, X = Y = 1, and the denominator of the R_o equation

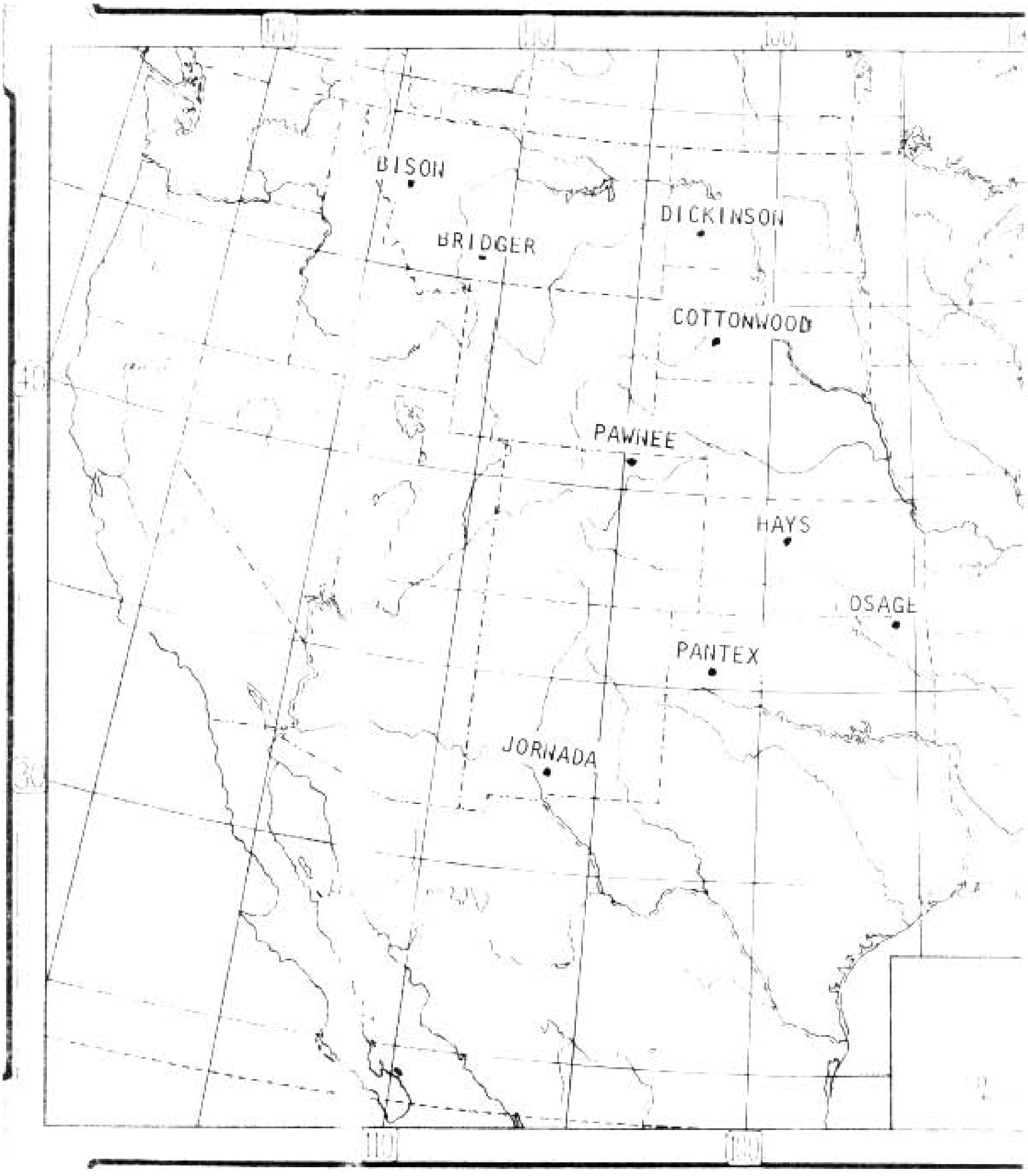


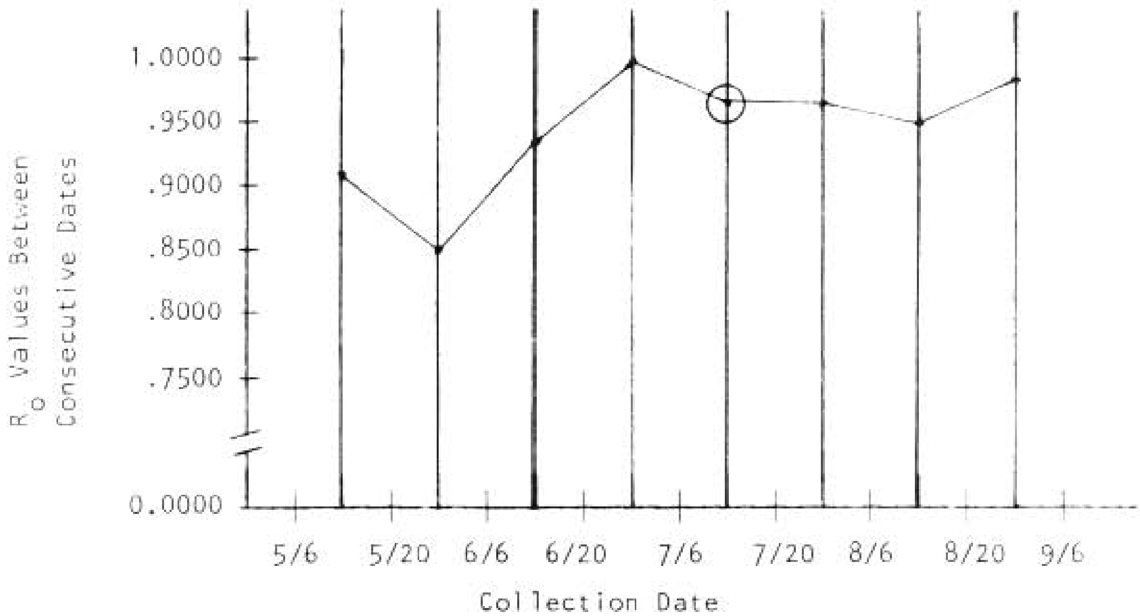
Fig. 1. Geographical location of the nine U.S. IBP Grassland Biome research sites compared in this report.

Table 1. Characteristics of the nine U.S. IBP Grassland Biome research sites compared in this report.

Site	Grassland Type	Major Vegetation	Elevation (m)	Annual Precipitation (cm)	Phenology (Lilac bloom)
Bison	Palouse	*Festuca scabrella* *Festuca idahoensis*	980	33	May 20
Bridger	Mountain	*Festuca idahoensis* *Agropyron subsecundum*	2,320	61	July 1
Cottonwood	Mixed	*Agropyron smithii* *Stipa viridula*	850	38	May 20
Dickinson	Mixed	*Stipa comata* *Bouteloua gracilis*	850	41	May 30
Hays	Mixed	*Andropogon scoparius* *Andropogon gerardi*	610	58	April 25
Jornada	Desert	*Bouteloua eriopoda* *Sporobolus flexuosus*	1,340	23	April 1
Osage	Tallgrass	*Andropogon gerardi* *Panicum virgatum*	380	94	April 20
Pantex	Shortgrass	*Bouteloua gracilis* *Aristida longiseta*	1,090	53	April 15
Pawnee	Shortgrass	*Bouteloua gracilis* *Buchloe dactyloides*	1,430	30	May 10

becomes the constant 1.3863. In the numerator, x_i and y_i represent the proportions of the samples x and y composed of species or group i.

In order to make intersite comparisons, it was necessary to select one sampling date from each site as representative. Plants collected on the various dates at the nine sites were grouped by species for comparisons between collection dates within each site. Each pair of x_i and y_i values represented the percentage of live aboveground biomass contributed by species i on two different dates. One collection date at each site was chosen to be used in comparisons between sites. The selection was made by computing the R_o values between consecutive collection dates within each site and plotting these values on a graph as shown in the example below (using data from Cottonwood).



A high R_o value indicates a high degree of similarity in the vegetation sampled on the two dates, and therefore little change between sampling periods. A collection date was then chosen from the "leveling off" portion of the graph which represented the period during the growing season when plant composition first attained a relatively stable condition (little change in dry weight species composition between successive collection dates). The date chosen for Cottonwood in this example was July 6, 1970. Table 2 lists the dates chosen for the other sites. Data from ungrazed treatment areas only were utilized in the intersite comparisons.

For intersite comparisons, it was necessary to combine taxa into larger plant groups in order to obtain sufficient overlap among sites for computation of meaningful R_o values. Few plant species were common to different sites, and no single species occurred at all sites. The final scheme that was adopted consisted of grouping all grasses into 7 tribes, all forbs into 24 families, and treating sedges and shrubs as 2 additional groups (see Appendix A). When this was done, each pair of x_i and y_i values represented the percentage of live aboveground biomass contributed by group i (either a grass tribe, a forb family, a sedge, or a shrub) at two different sites.

R_o values obtained for comparisons between sites were subjected to cluster analysis. A good review of clustering methods can be found on pages 305 to 312 of Sokal and Sneath's book, "Principles of Numerical Taxonomy" (1963). The method used in this analysis was the weighted pair-group method described below.

(i) All R_o values were arranged in a similarity matrix, part of which is reproduced below.

DICKINSON				
.5091	PAWNEE			
.3119	.2755	HAYS		
.1088	.0668	.7345	OSAGE	
.4658	.7129	.2072	.0449	JORNADA

Table 2. Collection dates used in comparisons between sites.

Site	Collection Date	Treatment
Bison	May 30, 1970	Treatment 1
Bridger	July 20, 1970	Ungrazed treatment, 4 ft snow fence)
Cottonwood	July 6, 1970	Treatment 1
Dickinson	June 22, 1970	Treatment 1
Hays	June 16, 1970	Treatment 1
Jornada	July 30, 1970	Treatment 1
Osage	July 16, 1970	Treatment 1
Pantex	July 13, 1970	Treatment 5
Pawnee	July 16, 1970	Ungrazed treatment

(ii) The pair of sites having the mutually highest R_o value were grouped together to form a cluster. For example, the R_o value comparing Hays and Osage was .7345. This was the highest R_o value for both Hays and Osage and, therefore, a mutually highest R_o value. Hays and Osage were then grouped together and new R_o values computed between this cluster (Hays plus Osage) and all other sites. The new R_o values were computed using Spearman's sums of variables formula which appears below (Sokal and Sneath 1963).

$$r_{qQ} = \frac{\square qQ}{\sqrt{q + 2\Delta q}\ \sqrt{Q + 2\Delta Q}}$$

r_{qQ} = the new correlation (R_o value) between group q and group Q

$\square qQ$ = the sum of all correlations between members of one group with the other group

Δq = the sum of all correlations between members of the first group

ΔQ = the sum of all correlations between members of the second group

q = the number of sites in the first group

Q = the number of sites in the second group

This formula yielded values which differed by no more than .05 from values computed directly from recombination of the original data into the appropriate clusters. It was also a much less tedious procedure. (Four different methods were used for recomputing R_o values. The weighted pair-group method using Spearman's sums of variables, as presented above, was judged to be the best in terms of accuracy and efficiency. See Appendix B for an outline of the other methods that were tried.)

(iii) A new similarity matrix was then constructed using these new R_o values as shown below.

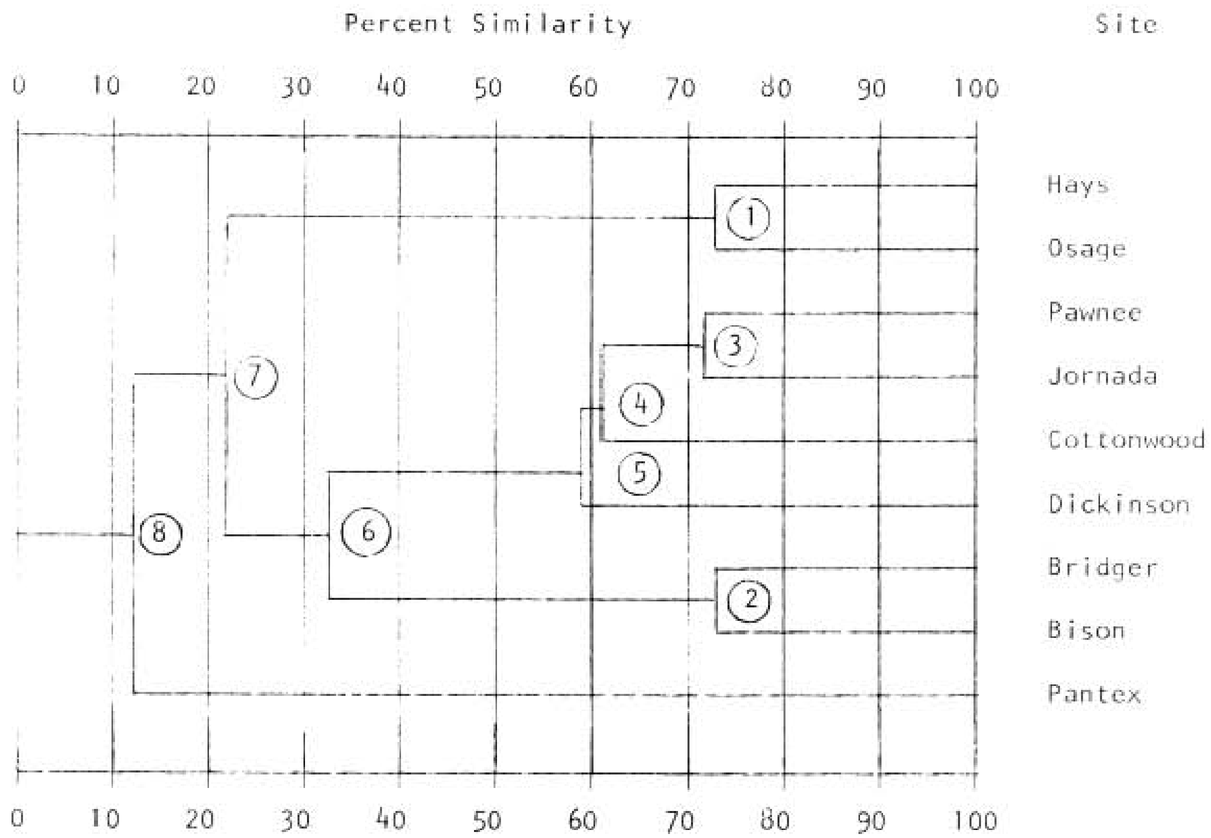
DICKINSON			
.5091	PAWNEE		
.2258	.1837	(HAYS + OSAGE)	
.4658	.7129	.1353	JORNADA

(iv) This procedure was repeated, each time grouping together the pair of sites or clusters having the mutually highest R_o value, until all sites were grouped into the same cluster. Note that for an R_o value to be mutually highest, it must be the highest for both sites concerned. For example, in the original similarity matrix, the highest R_o value for Dickinson was .5091 when compared with Pawnee. But Pawnee has a higher R_o value, .7129, when compared with Jornada which disqualifies the R_o value between Pawnee and Dickinson as being a mutually highest value.

A complete matrix of similarity indices from the original computation is shown in Table 3. A dendrogram was constructed representing the various R_o levels at which the sites were grouped together as a result of the cluster analysis (Fig. 2). Comparison of the similarity matrix of R_o values (Table 3) with the dendrogram can reveal an interesting insight into the mechanics of the cluster analysis. Note that in the similarity matrix Pawnee and Pantex have an R_o value of .6207, yet this relatively high correlation is completely masked in the dendrogram (Pantex remains isolated from the other sites until an R_o level of .1370). The reason for this seeming contradiction can be outlined as follows. The Pantex data used was composed of about 88%

Table 3. Similarity matrix of R_o values indicating floral similarities between sites. A value of 1 represents complete similarity; a value of 0 represents complete dissimilarity.

DICKINSON								
.2046	BISON							
.4299	.7314	BRIDGER						
.5455	.2267	.4762	COTTONWOOD					
.5091	.0891	.1199	.6254	PAWNEE				
.3119	.2648	.2674	.1683	.2755	HAYS			
.1088	.0073	.0066	.0478	.0663	.7345	OSAGE		
.0917	.0042	.0038	.2115	.6207	.1001	.0014	PANTEX	
.4658	.0279	.0168	.4939	.7129	.2072	.0449	.2261	JORNADA



Listed below are the R_o values at which the various sites were grouped together as a result of cluster analysis. $R_o \times 100$ = Percent similarity.

1.	(Hays) + (Osage)	.7345
2.	(Bison) + (Bridger)	.7314
3.	(Pawnee) + (Jornada)	.7129
4.	(Pawnee + Jornada) + (Cottonwood)	.6047
5.	(Pawnee + Jornada + Cottonwood) + (Dickinson)	.5985
6.	(Pawnee + Jornada + Cottonwood + Dickinson) + (Bison + Bridger)	.3315
7.	(Pawnee + Jornada + Cottonwood + Dickinson + Bison + Bridger) + (Hays + Osage)	.2290
8.	(Pawnee + Jornada + Cottonwood + Dickinson + Bison + Bridger + Hays + Osage) + (Pantex)	.1370

Fig. 2. Dendrogram representing percent similarity between sites as computed by cluster analysis of R_o values.

Opuntia sp. compared with about 21½% *Opuntia* sp. at Pawnee. It is this relatively large amount of cactus at Pawnee which is responsible for the .6207 R_o value with Pantex. When Pawnee is clustered with Jornada (at R_o = .7219) the *Opuntia* component of the cluster is essentially equal to the *Opuntia* component at Pawnee plus the *Opuntia* component at Jornada divided by two. But this, in effect, divides the Pawnee component by two because no *Opuntia* was collected at Jornada. Thus, the R_o value between Pantex and the cluster (Pawnee + Jornada) is reduced to .4496. During the next clustering cycle when the cluster (Pawnee + Jornada) is combined with Cottonwood, the same thing occurs, and the resulting R_o value between Pantex and the cluster (Pawnee + Jornada + Cottonwood) is .3690. At each successive clustering cycle the large *Opuntia* component at Pantex is compared with a smaller and smaller *Opuntia* component of the clustered sites because no *Opuntia* was collected at any other site except Pawnee. The result is successively lower R_o values between Pantex and the clustered sites until Pantex finally joins the cluster at R_o = .1370.

JACCARD'S COEFFICIENT OF COMMUNITY (J)

The equation is written below:

$$J = \frac{c}{a + b - c}$$

a = number of plant groups present at the first site

b = number of plant groups present at the second site

c = number of plant groups the two sites have in common

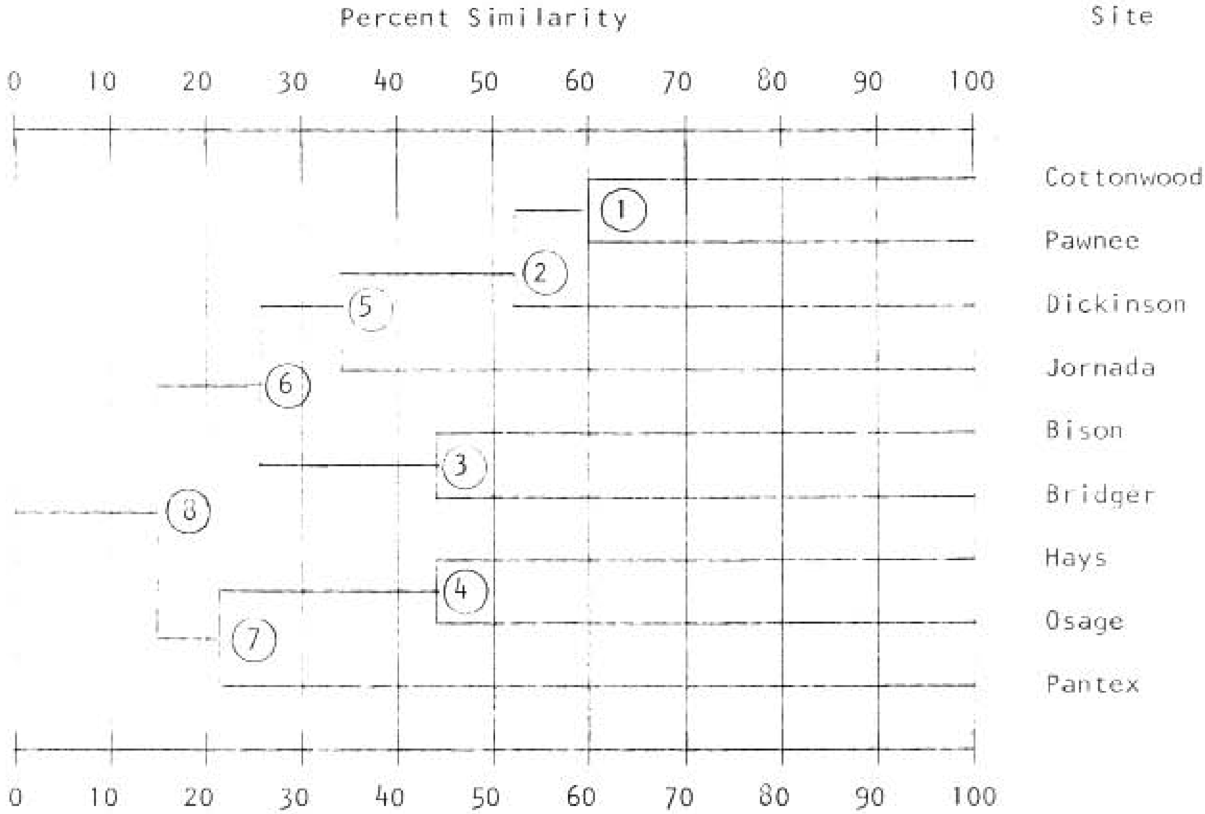
J can vary from zero (no plant groups in common) to one (all plant groups common to both sites).

Data from the same collection dates used for R_o comparisons between sites were used for J comparisons between sites (Table 2). The same plant groupings were also used (Appendix A). J values were computed between all sites using the equation written above. These values were then arranged in a similarity matrix (Table 4) and subjected to cluster analysis by the weighted pair-group method using Spearman's sums of variables to recompute J values at each stage of clustering (the same procedure that was used for the R_o cluster analysis). Fig. 3 is a dendrogram representing the various J levels at which the sites were grouped together as a result of the cluster analysis.

Note that Jaccard's coefficient of community is a measure of similarity which is based on the *number* of groups two sites have in common compared with the *number* of groups present at each site. In this respect it is an unweighted value, taking into account only group presence or absence and not differentiating between abundant and rare groups. The R_o equation does evaluate the proportional group composition, in terms of biomass, at each site when two sites are compared. This basic difference between the two indices should be kept in mind when comparing J values versus R_o values. Furthermore, these data were not collected to fulfill the requirements of this type of analysis. Plant species which contributed less than 5% of the total aboveground biomass were ignored. If the rare species had been included in the records, these would affect the values of J.

Table 4. Similarity matrix of J values indicating floral similarities between sites. A value of 1 indicates that all plant groups occurring at one also occur at the other site; a value of 0 represents sites with no plant groups in common.

DICKINSON								
.2500	BISON							
.2666	.4444	BRIDGER						
.4666	.2307	.2500	COTTONWOOD					
.4736	.2222	.2352	.6000	PAWNEE				
.3125	.1538	.1666	.2142	.3529	HAYS			
.2000	.0909	.1000	.1666	.1764	.4444	OSAGE		
.1428	.1111	.1250	.2000	.2000	.2222	.1428	PANTEX	
.3157	.2666	.2000	.3125	.3500	.2500	.1333	.1538	JORNADA



Listed below are the J values at which the various sites were grouped together as a result of cluster analysis. J × 100 = Percent similarity.

1.	(Cottonwood) + (Pawnee)	.6000
2.	(Cottonwood + Pawnee) + (Dickinson)	.5256
3.	(Bison) + (Bridger)	.4444
4.	(Hays) + (Osage)	.4444
5.	(Cottonwood + Pawnee + Dickinson) + (Jornada)	.3408
6.	(Cottonwood + Pawnee + Dickinson + Jornada) + (Bison + Bridger)	.2674
7.	(Hays + Osage) + (Pantex)	.2147
8.	(Cottonwood + Pawnee + Dickinson + Jornada + Bison + Bridger) + (Hays + Osage + Pantex)	.1506

Fig. 3. Dendrogram representing percent similarity between sites as computed by cluster analysis of J values.

ACKNOWLEDGMENTS

This report is based on field data collected by and generously made available by the following investigators: R. D. Pieper (JORNADA), R. D. Pettit (PANTEX), R. G. Risser (OSAGE), G. W. Tomanek (HAYS), P. L. Sims (PAWNEE), J. K. Lewis (COTTONWOOD), W. C. Whitman (DICKINSON), D. D. Collins (BRIDGLR), M. S. Morris (BISON). N. R. French gave advice on analytical methods and reviewed the manuscript.

LITERATURE CITED

Collins, D. 1970. Comprehensive Network Site description, BRIDGER. U.S. IBP Grassland Biome Tech. Rep. No. 1. 10 p.

Greig-Smith, P. 1964. Quantitative plant ecology. 2nd ed. Butterworth and Co., London. 256 p.

Herbel, C. H. and R. D. Pieper. 1970. Comprehensive Network Site description, JORNADA. U.S. IBP Grassland Biome Tech. Rep. No. 43. 21 p.

Horn, Henry S. 1966. Measurement of "overlap" in comparative ecological studies. Amer. Natur. 100:419-424.

Huddleston, E. W. 1970. Comprehensive Network Site description, PANTEX. U.S. IBP Grassland Biome Tech. Rep. No. 45. 12 p.

Jameson, D. A. and R. E. Bement. 1969. General description of the Pawnee Site. U.S. IBP Grassland Biome Tech. Rep. No. 1. 32 p.

Lewis, J. K. 1970. Comprehensive Network Site description, COTTONWOOD. U.S. IBP Grassland Biome Tech. Rep. No. 39. 26 p.

Morris, M. S. 1970. Comprehensive Network Site description, BISON. U.S. IBP Grassland Biome Tech. Rep. No. 37. 23 p.

Risser, P. G. 1970. Comprehensive Network Site description, OSAGE. U.S. IBP Grassland Biome Tech. Rep. No. 44. 5 p.

Shannon, C. E. and W. Weaver. 1949. The mathematical theory of communication. Univ. Illinois Press, Urbana. 125 p.

Sokal, R. R. and P. H. A. Sneath. 1963. Principles of numerical taxonomy. Freeman and Co., San Francisco. 359 p.

Tomanek, G. W. 1970. Comprehensive Network Site description, HAYS. U.S. IBP Grassland Biome Tech. Rep. No. 41. 6 p.

Whitman, W. C. 1970. Comprehensive Network Site description, DICKINSON. U.S. IBP Grassland Biome Tech. Rep. No. 40. 15 p.

APPENDIX A

This appendix contains a list of all plant species collected at the nine sites on the dates listed in Table 2. These species are presented in the groups to which they were assigned prior to intersite R_o calculations. The four letters preceding each species represent the computer code for that species used in data processing.

GRASSES

Tribe Agrostideae

ARLO *Aristida longesita* (Steud.)
ARPU *Aristida purpurea* (Nutt.)
CALO *Calamovilfa longifolia* (Hook.)
CAMO *Calamagrostis montanensis* (Scribn.)
MUTO *Muhlenbergia torreyi* (Kunth)
SPAS *Sporobolus asper* (Michx.)
SPFL *Sporobolus flexuosus* (Thurb.)
SPPI *Sporobolus pilosus* (Vasey.)
STCO *Stipa comata* (Trin.)
STVI *Stipa viridula* (Trin.)

Tribe Andropogoneae

ANGE *Andropogon gerardi* (Vitman.)
ANSC *Andropogon scoparius* (Michx.)
SONU *Sorghastrum nutans* (L.)

Tribe Aveneae

DAIN *Danthonia intermedia* (Vasey.)
KOCR *Koeleria cristata* (L.)

Tribe Chlorideae

BOCU *Bouteloua curtipendula* (Michx.)
BOER *Bouteloua eriopoda* (Torr).
BOGR *Bouteloua gracilis* (H.B.K.)
BUDA *Buchloe dactyloides* (Nutt.)

Tribe Festuceae

BRJA *Bromus japonicus* (Thunb.)
BRTE *Bromus tectorum* (L.)
ERPU *Erioneuron pulchellum* (Nash.)
FEID *Festuca idahoensis* (Elmer.)
FEOC *Festuca octoflora* (Walt.)
FESC *Festuca scabrella* (Torr.)
POPR *Poa pratensis* (L.)

Tribe Hordeae

AGSM *Agropyron smithii* (Rydb.)
AGSP *Agropyron spicatum* (Pursh)
AGSU *Agropyron subsecundum* (Link)
AGTR *Agropyron trachycaulum* (Link)
SIHY *Sitanion hystrix* (Nutt.)

Tribe Paniceae

PAVI *Panicum virgatum* (L.)

FORBS

Asclepiadaceae

ASVI *Asclepias viridiflora* (Raf.)

Boraginaceae

LARE *Lappula redowskii* (Hornem.)
LIRU *Lithospermum ruderale* (Dougl.)

Cactaceae

MAVI *Mammillari vivipara* (Nutt.)
OPFR *Opuntia fragilis* (Nutt.)
OPPO, OPU > *Opuntia polyacantha* (Haw.)

Caryophyllaceae

ARCO *Arenaria congesta* (Nutt.)
CEAR *Cerastium arvense* (L.)

Chenopodiaceae

ATAR *Atriplex argentea* (Nutt.)
ATCA *Atriplex canescens* (Pursh)
CHLE *Chenopodium leptophyllum* (Nutt.)
SAKA *Salsola kali* (L.)

Compositae

ACMI *Achillea millefolia* (L.)
AGGL *Agoseris glauca* (Pursh)
AMPS *Ambrosia psilostachya* (T.+G.)
ANMA *Anaphalis margaritacea* (L.)
ANRO *Antennaria rosea* (D. C. Eat.)
ARFU *Arnica fulgens* (Pursh)
ASAR *Aster arenosus* (Blake)
ASER *Aster ericoides* (L.)
ASFE *Aster fendleri* (Gray)
ASMU *Aster multiflorus* (Ait.)
ASOB *Aster oblongifolius* (Gray)
ASTA *Aster tanacetifolius* (H.B.K.)

Compositae (cont.)

BAOP *Bahia oppositifolia* (Nutt.)
CHVI *Chrysopsis villosa* (Pursh)
CIUN *Cirsium undulatum* (Nutt.)
ECAN *Echinacea angustifolia* (D.C.)
ERBE *Erigeron bellidiastrum* (Nutt.)
ERSP *Erigeron speciosus* (Lindl.)
HASP *Haplopappus spinulosus* (Greene)
HEPE *Helianthus petiolaris* (Nutt.)
HIAL *Hieracium albiflorum* (Hook.)
LAPU *Lactuca pulchella* (Pursh)
LASE *Lactuca serriola* (L.)
LIPU *Liatris punctata* (Hook.)
MINU *Microseris nutans* (Pursh)
RAT *Ratibida columnaris* (Sims.)
SECR *Senecio crocatus* (Rydb.)
SEPL *Senecio plattensis* (Nutt.)
SEUN *Senecio uintahensis* (A. Nels.)
SOMI *Solidago missouriensis* (Nutt.)
SOMO *Solidago mollis* (Bartl.)
SORI *Solidago rigida* (L.)
TEST *Tetraneuris stenophylla* (Rydb.)
THGR *Thelesperma gracile* (Torr.)
THME *Thelesperma megapotamicum* (Spreng.)
TOGR *Townsendia grandiflora* (Nutt.)
TRDU *Tragopogon dubius* (Scop.)

Cruciferae

DIWI *Dithyrea wislizeni* (Engelm.)
LEDE *Lepidium densiflorum* (Schrad.)

Euphorbiaceae

CRCO *Croton corymbulosus* (Klotzsch)

Labiatae

SCBR *Scutellaria brittonii* (Porter)
SCRE *Scutellaria resinosa* (L.)

FORBS (cont.)

Leguminosae

AMCA *Amorpha canescens* (Pursh)
ASST *Astragalus striatus* (Nutt.)
CABA *Cassia bauhinoides* (Michx.)
LOAM *Lotus americanus* (Nutt.)
LUAR *Lupinus argenteus* (Pursh)
LUSE *Lupinus sericeus* (Nutt.)
MEOF *Melilotus officinalis* (L.)
PEPU *Petalostemon purpureus* (Rydb.)
PSCU *Psoralea cuspidata* (Pursh)
PSES *Psoralea esculenta* (Pursh)
PSTE *Psoralea tenuiflorum* (Pursh)
SCUN *Schrankia uncinata* (Willd.)

Liliaceae

ALDR *Allium drummondii* (Regel)
ALTE *Allium textile* (Nels.)
FRPU *Fritillaria pudica* (Pursh)
YUEL *Yucca elata* (Nutt.)
YUGL *Yucca glauca* (Nutt.)
ZIPA *Zigadenus paniculatus* (Nutt.)

Linaceae

LIAU *Linum australe* (Heller)

Malvaceae

SPCO *Sphaeralcea coccinea* (Pursh)

Nyctaginaceae

MILI *Mirabilis linearis* (Pursh)

Onagraceae

GACO *Gaura coccinea* (Nutt.)
OECO *Oenothera coronopifolia* (T.+G.)
OENU *Oenothera nuttallii* (Sweet)
OESE *Oenothera serrulata* (Nutt.)
STIL *Stenosiphon linifolium* (Nutt.)

Orobanchaceae

ORLU *Orobanche ludoviciana* (Nutt.)

Plantaginaceae

PLPU *Plantago purshii* (Roem.)

Polemoniaceae

COLI *Collomia linearis* (Nutt.)

Polygonaceae

ERAB *Eriogonum abertianum* (Nutt.)
ERIO *Eriogonum* spp. (Michx.)

Primulaceae

DOCO *Dodecatheon conjugens* (Rydb.)

Rosaceae

GETR *Geum triflorum* (Pursh)

Saxifragaceae

HECY *Heuchera cylindrica* (Nutt.)
SARH *Saxifraga rhomboidea* (Greene)

Scrophulariaceae

CASU *Castilleja sulphurea* (Rydb.)

Violaceae

VINU *Viola nuttallii* (Pursh)

SEDGES

CAEL *Carex elynoides* (Holm)
CAFI *Carex filifolia* (Nutt.)
CAHE *Carex heliophila* (Mack.)
SEDE *Selaginella densa* (Rydb.)

SHRUBS

ARFR *Artemisia frigida* (Willd.)
ARLU *Artemisia ludoviciana* (Nutt.)
CHNA *Chrysothamnus nauseosus* (Pallas)
GUSA *Gutierrezia sarothrae* (Pursh)

APPENDIX B

This appendix briefly describes the four methods of cluster analysis to which the similarity matrix of R_o values (Table 3) was subjected. The goal was to find a clustering method that yielded results similar to those obtained when new similarity matrix coefficients were computed directly from recombination of the original data at each stage of clustering.

All four methods begin by choosing the pair of sites that have the highest mutual R_o value, combining those two sites to form a cluster, and then computing new R_o values between the newly formed cluster and all other sites. This same process is repeated at each successive clustering stage. The difference between the methods is the manner in which the new R_o values are computed. An example of the computation procedure is given for each method.

Method 1--R_o values computed by recombination of original data.

Hays and Osage were grouped together to form a cluster at R_o = .7345. The data from Hays and Osage was then combined as shown below.

Plant Group	Hays	Osage	(Hays + Osage)
Agrostideae	.012	.030	.021
Paniceae	.030	.029	.030
Andropogoneae	.576	.813	.694
Chlorideae	.079	.000	.040
Leguminosae	.182	.000	.091
Onagraceae	.024	.000	.012
Labiatae	.002	.000	.001
Compositae	.094	.002	.048
Sedges	.000	.014	.007
Unidentified Forbs	.000	.113	.056

The figures above represent percent composition by weight of the various plant groups. For example, Andropogoneae comprises 57.6% by weight of the total live aboveground plant biomass at Hays, 81.3% at Osage, and 69.4% when Hays and Osage are combined. New R_o values are then calculated by substituting the (Hays + Osage) percentages as the x_i's in the R_o equation:

$$R_o = \frac{\Sigma(x_i + y_i)\ \ln(x_i + y_i) - \Sigma\, x_i \ln x_i - \Sigma\, y_i \ln y_i}{(X + Y)\ \ln(X + Y) - X \ln X - Y \ln Y}$$

and computing new R_o values by letting the y_i's equal the corresponding percentages at each of the other sites in turn. The R_o levels at which the various sites clustered together when this method was used are shown below.

1.	(Hays) + (Osage)	.7345
2.	(Bison) + (Bridger)	.7314
3.	(Pawnee) + (Jornada)	.7129
4.	(Pawnee + Jornada) + (Cottonwood)	.5844
5.	(Pawnee + Jornada + Cottonwood) + (Dickinson)	.6169
6.	(Pawnee + Jornada + Cottonwood + Dickinson) + (Bison + Bridger)	.3789
7.	(Pawnee + Jornada + Cottonwood + Dickinson + Bison + Bridger) + (Hays + Osage)	.2743
8.	(Pawnee + Jornada + Cottonwood + Dickinson + Bison + Bridger + Hays + Osage) + (Pantex)	.1325

Method 2--R_o values computed by arithmetic mean.

Hays and Osage were grouped together to form a cluster at R_o = .7345. The arithmetic mean of the old R_o value between Hays and a given site and the old R_o value between Osage and the same site was then calculated. This mean value became the new R_o value between the cluster (Hays + Osage) and the given site. The R_o levels at which the various sites clustered together when this method was used are shown below.

1.	(Hays) + (Osage)	.7345
2.	(Bison) + (Bridger)	.7314
3.	(Pawnee) + (Jornada)	.7129
4.	(Pawnee + Jornada) + (Cottonwood)	.5596
5.	(Pawnee + Jornada + Cottonwood) + (Dickinson)	.5164
6.	(Pawnee + Jornada + Cottonwood + Dickinson) + (Bison + Bridger)	.2623
7.	(Pawnee + Jornada + Cottonwood + Dickinson + Bison + Bridger) + (Hays + Osage)	.1529
8.	(Pawnee + Jornada + Cottonwood + Dickinson + Bison + Bridger + Hays + Osage) + (Pantex)	.0776

Method 3--R_o values computed by Spearman's sums of variables, unweighted pair-group.

Hays and Osage were grouped together to form a cluster at R_o = .7345. The new R_o values were then computed from Spearman's sums of variables formula written below:

$$r_{qQ} = \frac{\square qQ}{\sqrt{q + 2\Delta q}\ \sqrt{Q + 2\Delta Q}}$$

r_{qQ} = the new correlation (R_o value) between group q and group Q

$\square qQ$ = the sum of all correlations between members of one group with the other group

Δq = the sum of all correlations between members of the first group

ΔQ = the sum of all correlations between members of the second group

q = the number of sites in the first group

Q = the number of sites in the second group

In the unweighted version, the new R_o values at each clustering stage are computed via the above formula from the R_o values in the *original* similarity matrix rather than from the immediately preceding matrix (this differs from the weighted version shown in method 4).

Assume Pawnee and Jornada are a newly formed cluster. The new R_o value between this new cluster and the already clustered group (Bison + Bridger) is calculated in the following manner.

R_o values from original similarity matrix.

BISON			
.7314	BRIDGER		
.0891	.1199	PAWNEE	
.0279	.0168	.7129	JORNADA

$\square qQ$ = (R_o between Bison and Pawnee) + (R_o between Bison and Jornada) + (R_o between Bridger and Pawnee) + (R_o between Bridger and Jornada)

$= (.0891) + (.0279) + (.1199) + (.0168) = .2537$

Δq = (R_o between Pawnee and Jornada) = .7129

ΔQ = (R_o between Bison and Bridger) = .7314

$q = 2$

$Q = 2$

$$r_{qQ} = \frac{\square qQ}{\sqrt{q + 2\Delta q}\ \sqrt{Q + 2\Delta Q}} = \frac{.2537}{\sqrt{2 + (2)(.7129)}\ \sqrt{2 + (2)(.7314)}}$$

$= .0736$ = the new R_o value between (Pawnee + Jornada) and (Bison + Bridger)

The R_o levels at which the various sites clustered together when this method was used are shown below.

1.	(Hays) + (Osage)	.7345
2.	(Bison) + (Bridger)	.7314
3.	(Pawnee) + (Jornada)	.7129
4.	(Pawnee + Jornada) + (Cottonwood)	.5596
5.	(Pawnee + Jornada + Cottonwood) + (Dickinson)	.5889*
6.	(Pawnee + Jornada + Cottonwood + Dickinson) + (Pantex)	.3514
7.	(Pawnee + Jornada + Cottonwood + Dickinson + Pantex) + (Bison + Bridger)	.2296
8.	(Pawnee + Jornada + Cottonwood + Dickinson + Pantex + Bison + Bridger) + (Hays + Osage)	.2221

*This seeming inconsistency (.5889 being higher than the .5596 R_o level of the previous clustering stage) is due to the mechanics of the computations involved. It is not uncommon when using Spearman's formula to find a site that has a slightly higher R_o value with a newly formed cluster than it did with either component of the cluster before they were joined together.

Method 4--R_o values computed by Spearman's sums of variables, weighted pair-group.

Hays and Osage were grouped together to form a cluster at R_o = .7345. The new R_o values were then computed from Spearman's sums of variables formula (same formula as in Method 3 above). The mechanics are the same as in the unweighted method except that it is the immediately preceding matrix from which values are taken to compute the new R_o values. The same example that was used in Method 3 is shown again below, this time demonstrating the form of the calculations when the weighted method is employed.

Assume Pawnee and Jornada are a newly formed cluster. The new R_o value between the new cluster and the already clustered group (Bison + Bridger) is calculated in the following manner.

R_o values from preceding similary matrix.

BISON + BRIDGER		
.1123	PAWNEE	
.0240	.7129	JORNADA

$\square qQ$ = [R_o between (Bison + Bridger) and Pawnee] + [R_o between (Bison + Bridger) and Jornada] = (.1123) + (.0240) = .1363

Δq = (R_o between Pawnee and Jornada) = .7129

ΔQ = 0, Bison and Bridger are treated as a single entity, and therefore have no R_o value between them (Note: This is not an R_o value of zero representing complete dissimilarity between sites, but rather the zero represents the absence of any R_o value at all).

$q = 2$

$Q = 1$, Bison and Bridger are treated as a single entity.

$$r_{qQ} = \frac{\square qQ}{\sqrt{q + 2\Delta q}\ \sqrt{Q + 2\Delta Q}} = \frac{.1363}{\sqrt{2 + (2)(.7129)}\ \sqrt{1 + (2)(0)}}$$

= .0736 = the new R_o value between (Pawnee + Jornada) and (Bison + Bridger)

Note that in this example the new R_o value obtained by the weighted version of the method is the same as the new R_o value obtained by the unweighted version presented in method 3. This is not usually the case, although the two methods do yield quite similar results.

The R_o levels at which the various sites clustered together when the weighted method was used are shown below.

1.	(Hays) + (Osage)	.7345
2.	(Bison) + (Bridger)	.7314
3.	(Pawnee) + (Jornada)	.7129
4.	(Pawnee + Jornada) + (Cottonwood)	.6047
5.	(Pawnee + Jornada + Cottonwood) + (Dickinson)	.5985
6.	(Pawnee + Jornada + Cottonwood + Dickinson) + (Bison + Bridger)	.3315
7.	(Pawnee + Jornada + Cottonwood + Dickinson + Bison + Bridger) + (Hays + Osage)	.2290
8.	(Pawnee + Jornada + Cottonwood + Dickinson + Bison + Bridger + Hays + Osage) + (Pantex)	.1370

Comparison of Methods

Values obtained by Method 4 (Spearman's sums of variables, weighted pair-groups) were found to more closely approximate the values obtained by Method 1 (recombination of original data) than did either of the other methods. Method 4, in addition to approximating values obtained from recombination of the original data, also proved much less tedious, and therefore more efficient when a fairly large amount of data must be handled.

Listed below are the average differences and maximum differences in R_o values between Method 1 and each of the other methods.

Method	Average Difference	Maximum Difference
Method 2	.0522	.1214
Method 3	.0268	.0896
Method 4	.0169	.0474