

DEVELOPMENT OF THE UPPER LEVEL SYNOPTIC INDEX

K.L. Frank¹, B.W. Geils², L.S. Kalkstein¹, H. Thistle³

Draft – 1 March 2005

¹ Center for Climatic Research, Department of Geography, University of Delaware, Newark, DE 19716. e-mail: klfrank@udel.edu, phone: (302) 831-2254, fax: (302) 831-6654. ² USDA Forest Service, Rocky Mountain Research Station, Flagstaff, AZ 86001. ³ USDA Forest Service, Forest Health Technology Enterprise Team, Morgantown, WV 26505.

Abstract

This study developed a methodology to temporally classify large scale, upper level atmospheric conditions over North America, utilizing the newly-developed Upper Level Synoptic Index (ULSI). Four meteorological variables: geopotential height, specific humidity, and u- and v-wind components, at the 500 mb level over North America were obtained from the NCEP/NCAR Reanalysis Project data set for the period 1965–1974. These data were subjected to principal components analysis to standardize and reduce the data set, and then an average linkage clustering algorithm identified groups of observations with similar flow patterns. The procedure yielded 16 clusters. These flow patterns identified by the ULSI typify all patterns that are anticipated to be observed over the study area. Additionally, the resulting cluster calendar for the period 1965–1974 showed that the clusters are generally temporally continuous. The ULSI calendar of synoptic conditions can be used to identify situations that lead to periods of extreme weather and to explore transport of airborne particles and biota across North America.

Keywords

Synoptic, Indexing, Classification, North America

Introduction

The term *synoptic climatology* stems from the Greek word *sunopsis*, or general view (American Heritage Dictionary, 1991). Synoptic climatology presents a holistic view of the climate system, instead of focusing on the individual elements, or variables, that combine to make up the system. This perspective is especially applicable to investigation of how biological systems interact with the climate system because organisms do not respond to changes in single weather variables; rather, they respond to the combination of variables working simultaneously to make up the overall climate condition.

Classification, or indexing, is often employed in synoptic climatological analyses because it provides a simple way to summarize the combinations of variables working together at a given time. Indices are abstractions that permit us to categorize an immense mass of data. Indexing emphasizes the differences between meteorological events that are useful to solve environmental problems. This study outlines the development of the Upper Level Synoptic Index (ULSI), an index that categorizes upper level flow patterns across North America in order to address certain environmental issues.

Synoptic Climate Indices

A synoptic climate index classifies overall weather conditions into groups with minimal within group variance and maximal between group variance (Balling, 1984). Consequently, a synoptic classification attempts to accentuate meteorological differences among categories. Indexing should also minimize the differences within each class. Yarnal (1993) identifies two major types of classification methods: manual and automated.

In general, manual classification methods entail knowledge-based visual interpretation of individual weather maps. These classification schemes are subjective, time and labor intensive, and generally not repeatable (Sheridan, 2002). Applications of manual classification schemes include forecasting (Elliott, 1949); mesoscale modeling (Pielke and Segal, 1986); environmental analysis, such as variation in pollution concentrations (Davies *et al.*, 1986, Muller and Jackson, 1985); and bioclimatological investigations (Muller, 1985).

Automated classification schemes employ statistical algorithms, so are generally fast to construct and easily replicated. Some subjectivity is introduced into these otherwise objective clustering schemes through the selection of variables. The analyst should choose only those variables thought to be important for the intended application.

Development of automated synoptic indices using cluster analytic techniques began in the mid-1970s as computing capacity enabled researchers to employ computationally intensive algorithms on larger and larger data sets. Ayoade (1977), for example, performed a climate regionalization for Nigeria by applying several clustering algorithms to a data matrix of 10 meteorological variables from 32 stations.

Kalkstein and Corrigan (1986) developed the Temporal Synoptic Index (TSI), a calendar identifying days with similar overall surface conditions at a single location. The TSI does not account for upper level conditions; but they may be induced by changing the input data matrix. A more fundamental limitation is that, as a point index, TSI does not recognize the spatial continuity of air masses. Therefore, when the same clustering procedure was applied to proximate stations the resulting clusters were less similar than desired. This limited use of the TSI to applications considering only single locations.

The Spatial Synoptic Index (Davis and Kalkstein, 1990a) expanded the TSI into a spatial index that "...is capable of dividing a large region into homogeneous synoptic categories" (Davis and Kalkstein, 1990b, p. 322). Although the Spatial Synoptic Index is spatially continuous, it is temporally discontinuous; that is, air masses are similar at nearby locations but, over time, air masses arise from no apparent source.

To account for both spatial and temporal continuity of synoptic conditions, Vose (1993) developed the Regional Synoptic Index (RSI). This index considered a matrix of six variables at seven stations across a region and classified the observations following the method of the TSI. The result is a temporal index where the synoptic categories are expressed as regional maps.

Integration of upper level data to cluster-based synoptic indexing began with Webber's (1994) Upper Air Synoptic Classification (UASC). The UASC applied the techniques of the TSI to gridded upper level observations across North America. Because of the massive size of the study area, only two variables (observed once daily at 459 points for three levels) were included in the development of the UASC. This index showed that the technique was able "to create a rational classification of large-scale upper air circulation" (Webber, 1994, p. 102). The applicability of the UASC is limited, however, because it produced excessively inclusive clusters. For example, the UASC resulted in only one summer cluster; representation of all days in a season within a single cluster seriously restricts its utility in many environmental studies. This problem may be alleviated by increasing the number of clusters allowed by the algorithm, however the inclusion of only two variables for analysis may continue to restrict the viability of this classification.

Schreiber (1996) incorporated both surface and upper level data in the development of the Surface/Upper Level Synoptic Index (SULSI). This index employed the techniques of the TSI and considered four-times-daily surface observations of six variables and twice-daily observations of five variables at three heights aloft. A small study area in the Desert Southwest allowed the incorporation of so many variables but restricted the geographic extent to which it could be employed.

The Upper Level Synoptic Index (ULSI) presented here identifies similar upper level synoptic situations from numerous observations of many variables (without surface observations). The ULSI improves upon the UASC which considers only temperature at the pressure surface height by considering both humidity and wind at the pressure surface.

Upper Level Synoptic Index

The development of the ULSI begins with four-times-daily data for 4 meteorological variables at the 500 mb level available from the NCEP/NCAR Reanalysis Project data set (Kalnay *et al.*, 1996, NOAA-CIRES, 2002). This data set is generated from the application of mathematical data assimilation and forecasting models to historical weather data from numerous sources (UCAR, 2003). The result is a global, gridded, four-times-daily data set with a period of record beginning January 1, 1948. This data set is especially useful for application to upper level synoptic indexing; it provides a much finer spatial resolution (2.5°) than raw radiosonde sounding data and is standardized to account for changes in observational techniques and missing observations.

To accommodate anticipated and potential applications, the ULSI is developed for a large region and long period of record. The study region extends from 20°N to 60°N and 60°W to 140°W including all of the conterminous United States and much of Mexico and Canada (Figure 1). The initial period of record indexed is from 1965 to 1974. This period was chosen based on the desire to apply the Index to a specific event during the period. Additional days are indexed by comparing the conditions on a given day to the mean conditions of each of the clusters and placing an observation into the existing cluster to which the observation is most similar (using an equally weighted sum of squared \bar{x} -scores, see Sheridan, 2002).

The data matrix used to determine the ULSI includes four variables observed at the 500 mb level four times a day on a regular spatial grid (Table 1). The 500 mb level was selected for this analysis because it is generally midway between the levels of divergence and non-divergence; this level is commonly thought of as representative of the steering circulation for surface systems. The variables representing conditions at the 500 mb level are height of the pressure surface (m), specific humidity (kg/kg), u-wind component (m/s), and v-wind component (m/s). Air temperature was excluded from this analysis because it is highly correlated with the height of the 500 mb surface. The original grid of 2.5° latitude by 2.5° longitude is reduced to 3.54° resolution by eliminating every other point; results will show that this data reduction does not affect the outcome of the classification in a negative manner.

The data matrix was subjected to an average linkage clustering algorithm (Sokal and Michener, 1958). Kalkstein *et al.* (1987) found this clustering method to be most appropriate for environmental analyses.

Introduction of raw weather data to the clustering algorithm gives each variable presented to the algorithm equal weight. This is problematic if the variables introduced have different units or if any of them are collinear. To avoid these problems, principal components analysis (PCA) was applied to the original data set before the clustering procedure began.

The 87 components with variance greater than 1.0 were retained and input to the clustering algorithm (Kaiser, 1960). These components explained 96.6% of the variance within the original data set. In this case, the first two components loaded most heavily on geopotential height in the northwest corner of the study area and explained over 32% of the variance in the data set. The third and fourth components had heavier loadings on the wind and humidity variables and were more strongly associated with values from the central part of the study area. The first five components explained over half of the variance in the original data.

Eleven main clusters resulted from the input of the components to the clustering algorithm. In an effort to identify potentially important within-cluster variations that could be present in clusters containing a large number of observations, any cluster that contained more than 20% of the total observations was then resubmitted to the procedure to yield a set of “nested” clusters. A total of 16 clusters were identified (Table 2). For illustrative purposes, a sum of squared χ -scores method was used to identify the observation most similar to the cluster mean and the 500 mb height contours plotted for each cluster.

Results

As is expected for upper level flow patterns, examination of the resulting cluster calendar for the period 1965–1974 shows that the clusters are generally temporally continuous. Nearly 55% of all observations over the 10-year period are preceded by an observation in the same cluster.

Only the three least frequently occurring clusters are not most often preceded by themselves (Table 3).

Winter Clusters

Six clusters occur predominantly in the winter months (December, January, February) (Figure 2). Winter clusters, as expected, are generally characterized by lower 500 mb heights in the northern part of the study region with these low heights extending considerably south under some situations. The winter clusters generally exhibit more zonal flow than any other season, although meridional flow is represented by some clusters. Almost one-third of all observations fall in clusters with winter frequency maxima.

Observations falling in cluster 1 (Winter-Trough-Ridge) are associated with a short-wave trough in the western United States and a ridge in the eastern United States. Flow across the southern part of the study region is somewhat meridional, but more zonal conditions prevail in the north. This cluster exhibits fairly strong baroclinicity, indicated by a fairly steep height gradient. This pattern occurs most frequently in December and January although it is observed in all months of the year.

Conversely, Cluster 3.2 (Winter-Ridge-Trough) observations denote a ridge in the west and a trough in the eastern United States. In this case, lower 500 mb heights are observed much farther south and the zone of baroclinicity is much more pronounced and encompasses nearly the entire study area. This cluster exhibits the strongest baroclinic flow of all clusters. Winter-Ridge-Trough patterns occur in all months but are most frequent in January, February and March.

Observations with Winter-Zonal-Barotropic flow are identified in cluster 4. Two low-amplitude short-wave features are present in the northern part of the study area—a ridge in the west and an inverted trough in the east. The degree of baroclinicity is less under this situation than in cluster 1 or 3.2. Winter-Zonal-Barotropic conditions occur most frequently in December, but cluster 4 is more frequent in the summer and transition months than any other cluster with a winter frequency maximum.

Cluster 7 (Inverted Trough) portrays an inverted trough in the central U.S. Inverted Trough shows the strongest meridional flow of the winter clusters (particularly in the West) with the trough axis along a line from the Great Lakes to Baja California. This condition occurs less than 2% of the time and is observed most often in January. Cluster 7 does not occur between May and September.

A particularly zonal upper level pattern is identified by cluster 8 (Winter-Zonal-Baroclinic). These observations display lower 500 mb heights farther south and a stronger zone of baroclinicity than those observations classed as Winter-Zonal-Barotropic (Cluster 4). Winter-Zonal-Baroclinic conditions are most frequent in January and are not observed at all between April and September.

Cluster 11 is similar to cluster 3.2 in that it is a Winter-Ridge-Trough pattern. However, the pattern in cluster 11 is shifted north from that seen in cluster 3.2 and the flow is considerably less baroclinic. A particularly strong ridge is evident over the southwestern United States. Winter-Ridge-Trough (northerly displacement) observations occur very infrequently, only 0.04% of the time. These conditions are present only in December, January and February and are most frequent in January.

Transition Season Clusters

Three clusters are observed primarily in the transition months (March-May and September-November) (Figure 3). These clusters are characterized by a general meridional flow, especially in the southwestern United States. In addition, these three categories are similar in that they depict fairly strong baroclinic conditions. These clusters represent the strong latitudinal contrast between cold and warm air, expected during the transition months. Less than 20% of all observations are placed in clusters with transition season maxima.

The most frequent transition season cluster is 3.1 (Transition-Trough-Ridge), which is characterized by a trough in the western United States and a ridge in the eastern part of the country. Conditions are generally baroclinic, especially in the western reaches of the region. This baroclinic zone is stronger and shifted southward and westward from the pattern observed in Winter-Trough-Ridge (cluster 1), a category it resembles. Cluster 3.1 is most frequent in March and October but is observed in all months.

Observations with a high amplitude trough in the central and ridge in the western United States are grouped in cluster 3.3 (Transition-Ridge-Trough (high amplitude)). This cluster displays the greatest wave amplitude of all clusters, and the ridge over the West is particularly well-pronounced. The wave axes are slightly inverted in cluster 3.3 but not as strongly as those in cluster 7 (Inverted Trough). Transition-Ridge-Trough (high amplitude) conditions occur most frequently in June and September but are present in all months.

A second group of much lower amplitude Transition-Ridge-Trough observations comprises cluster 2.4. The zone of baroclinicity is shifted north from that in cluster 3.3 and the degree of

zonal flow in the southern reaches of the area is much higher. The wave features in Transition-Ridge-Trough have nearly the same amplitude as those of cluster 3.2 (Winter-Ridge-Trough) but are shifted slightly to the west. Transition-Ridge-Trough occurs in all months but makes up less than one percent of the total observations. Cluster 2.4 occurs most frequently in the spring.

Summer Clusters

Seven clusters occur most frequently in the summer months (June, July, August) (Figure 4). Summer clusters are more barotropic than other seasons, as would be expected when the zone of baroclinicity shifts farther north and weakens. Summer clusters are also characterized by higher 500 mb heights that extend farther north into the study area, indicating the presence of warmer air. More than half of all observations were identified in clusters with summer maxima.

Cluster 2.1 (Summer-Trough-Ridge) is characterized by the same general flow pattern as cluster 1 (Winter-Trough-Ridge), but the height of the 500 mb surface is much higher in the former. The amplitude of the waves is similar in the summer and winter Trough-Ridge clusters. The waves in Summer-Trough-Ridge are shifted slightly north from those in Winter-Trough-Ridge and are much farther north than Transition-Trough-Ridge. A warm cored anticyclone is obviously present over the East, which is a common circulation feature in the summer. Cluster 2.1 is the most frequently occurring cluster (15.22%) year-round and is found in all months. Summer-Trough-Ridge is most abundant in July and August.

A trough in the western United States is also representative of cluster 2.2 (Summer-Trough), but the 500 mb heights are much higher over the entire country and the main zone of

baroclinicity is almost entirely in Canada. No distinct ridge is present in the eastern United States as it is in the Trough-Ridge clusters. The amplitude of the trough in the west is greater than that of cluster 1 and cluster 2.1 and the degree of baroclinicity is stronger around the trough and extends to higher 500 mb heights. Cluster 2.2 is most frequent in the period May-August, but is present in all months.

A Trough-Ridge-Trough flow is present in cluster 2.3 although the eastern trough is much less well-defined than that in the west where the heights are much lower than those of either cluster 2.1 or cluster 2.2. Strong baroclinicity extends south to nearly 30°N and encompasses higher 500 mb heights than cluster 2.1 or 2.2. The baroclinic zone retreats to higher latitudes in the eastern part of the study region. Cluster 2.3 generally represents a shift north and west of the pattern observed in cluster 3.1. This pattern is present year-round but is most often observed in June and September.

Observations with a zonal flow and a northerly displacement of the jet are grouped in cluster 5, Summer-Zonal. This most zonal of the summer clusters exhibits a higher degree of meridional flow and a stronger zone of baroclinicity than the winter zonal clusters (clusters 4 and 8). Like many summer clusters, Summer-Zonal exhibits a warm-cored anticyclone in the Southeast. This pattern occurs most frequently in July and August but is present year-round.

Cluster 6 is characterized by a trough-ridge-trough pattern but is distinguished from cluster 2.3 by the northerly displacement of the zone of baroclinicity such that much of the United States is under barotropic flow. This cluster shows the greatest northern displacement of the highest 500 mb heights. The very high 500 mb heights across much of the southern United States

indicate very warm conditions across that part of the study region. These conditions occur most often in August but are present to some degree in all months.

Clusters 9 and 10 both exemplify a high amplitude trough-ridge-trough flow across North America. Both are characterized by high 500 mb heights extending very far north in the Plains, surrounded by regions of low heights in the northwest and northeastern corners of the United States. These areas of lower heights are still higher than levels observed at the same latitudes in many of the winter clusters. The barotropic ridge is more zonal than the other summer clusters. Clusters 9 and 10 together comprise only 0.36% of all observations and are most often observed in June and July.

Discussion

The ULSI successfully identifies observations with similar upper level flow patterns across North America and places these observations into clusters of reasonable size and character. The resulting clusters have high within-group similarity and high between-group dissimilarity. Examination of the resulting cluster calendar for the period 1965–1974 shows that the clusters are generally temporally continuous, as would be expected, and the 16 flow patterns identified by the ULSI exemplify all patterns that are anticipated to be observed over the study area. When compared to the previous synoptic typing of summertime upper level flow in the Desert Southwest completed by Carleton (1987), the summer ULSI clusters appear representative of the synoptic types previously identified. This indicates that the clusters identified by the ULSI are, indeed, representative of the types of upper level flow observed in the region.

The ULSI clusters presented here provide a basis for classification of additional observations. Using the mean conditions for each cluster, new observations may be classified by comparing

them to the clusters through the application of an equally weighted sum of squared z -scores method, similar to that outlined by Sheridan (2002). This will enable the production of a calendar of upper level patterns for the full period of record of the data set, 1948 to present.

Scores of applications present themselves to such a calendar of upper level patterns. This methodology may be useful in understanding movements of forest and agricultural pathogens and aerosols. The ULSI calendar is potentially effective for identifying synoptic situations that lead to periods of flood or drought in North America. On a somewhat shorter temporal scale, the ULSI could be applied to identify flow patterns that result in heat or cold waves. A temporal evaluation of the ULSI could help in the understanding of climate change issues. For example, identical flow patterns could today lead to more severe or extended heat waves or longer lasting wet or dry periods than they did in the past. Finally, coupling of the ULSI with surface air mass calendars, like the SSC2 (Sheridan, 2002), will allow investigation of the synoptic flow conditions that precipitate surface situations that have been identified as harmful to human health.

ACKNOWLEDGMENTS

Funded by USDA Forest Service as Cooperative Agreement 01-CA-11244225-231 (Forest Health Technology Enterprise Team) and Research Joint Venture 03-JV-11221605-297 (Rocky Mountain Research Station).

REFERENCES

- American Heritage Dictionary. 1991. *American Heritage Dictionary, Second College Edition*. Houghton Mifflin Company: Boston.
- Ayoade JO. 1977. On the use of multivariate techniques in climatic classification and regionalization. *Archiv fur Meteorologie Geophysik und Bioklimatologie, Series B*. **24**: 257-267.
- Balling RC, Jr. 1984. Classification in climatology. In *Spatial Statistics and Models*, Gaile GL, Willmott CJ (eds). D. Reidel Publishing Co.: Hingham, Massachusetts; 81-108.
- Carleton AM. 1987. Summer circulation climate of the American Southwest, 1945 – 1984. *Annals of the Association of American Geographers*. **77**(4):619-634.
- Davies TD, Kelly PM, Brimblecombe P, Farmer G, Barthelmie RJ. 1986. Acidity of Scottish rainfall influenced by climatic change. *Nature*. **322**: 359-361.
- Davis RC, Kalkstein LS. 1990a. Development of an automated spatial synoptic climatological classification. *International Journal of Climatology*. **10**: 769-794.
- Davis RC, Kalkstein LS. 1990b. Using a spatial synoptic climatological classification to assess changes in atmospheric pollution concentrations. *Physical Geography*. **11**(4): 320-342.
- Elliott RW. 1949. The weather types of North America. *Weatherwise*. **2**: 15-18, 40-43, 64-67, 86-88, 110-113, 136-138.
- Kaiser HF. 1960. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*. **20**: 141-151.
- Kalkstein LS, Corrigan P. 1986. A synoptic climatological approach for geographical analysis: assessment of sulfur dioxide concentrations. *Annals of the Association of American Geographers*. **76**(3): 381-395.

- Kalkstein LS, Tan G, Skindlov JA. 1987. An evaluation of three clustering procedures for use in synoptic climatological classification. *Journal of Climate and Applied Meteorology*. **26**(6): 717-730.
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR reanalysis 40-year project. *Bulletin of the American Meteorological Society*. **77**: 437-471.
- Muller RA. 1985. The potential for the atmospheric transport of moths from the perspective of synoptic climatology. In *Movement and Dispersal of Environmentally Important Biotic Agents: an International Conference on the Movement and Dispersal of Biotic Agents: A Symposium Held at Baton Rouge, Louisiana, October 17-19, 1984*. Claitor's Publishing Division: Baton Rouge, Louisiana; 179-202.
- Muller RA, Jackson AL. 1985. Estimates of climatic air quality potential for Shreveport, Louisiana. *Journal of Climate and Applied Meteorology*. **24**: 293-301.
- NOAA-CIRES Climate Diagnostics Center. 2002.
<http://www.cdc.noaa.gov/cdc/data.ncep.reanalysis.html>. [Accessed March 12, 2002].
- Pielke RA, Segal M. 1986. Mesoscale circulations forced by differential terrain heating. In *Mesoscale Meteorology and Forecasting*, Ray PS (ed). American Meteorological Society: Boston; 516-548.
- Schreiber KV. 1996. *A Synoptic Climatological Approach to Assessment of Visibility and Pollutant Source Locations, Grand Canyon National Park Area*. Ph.D. Dissertation, Department of Geography, University of Delaware: Newark, Delaware.

- Sheridan SC. 2002. The redevelopment of a weather-type classification scheme for North America. *International Journal of Climatology*. **22**: 51-68.
- Sokal RR, Michener CD. 1958. A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*. **38**: 1409-1438.
- UCAR. April 1, 2003. *NCEP/NCAR Reanalysis > Project Description*.
http://dss.ucar.edu/pub/reanalysis/rean_proj_des.html. [accessed August 4, 2003].
- Vose RS. 1993. *Development of a Regional Synoptic Index for Environmental Analysis*. M.S. Thesis, Department of Geography, University of Delaware: Newark, Delaware.
- Webber SR. 1994. *Development of a Synoptic Classification of North American Atmospheric Circulation*. M.S. Thesis, Department of Geography, University of Delaware: Newark, Delaware.
- Yarnal B. 1993. *Synoptic Climatology in Environmental Analysis*. Belhaven Press: London.

Table 2. Characteristics of the 16 clusters identified by the Upper Level Synoptic Index.

Cluster Number	Description	Season	Frequency (%)
1	Trough-Ridge	Winter	10.31
2*			33.58
2.1	Trough-Ridge	Summer	15.22
2.2	Trough	Summer	13.01
2.3	Trough-Ridge-Trough	Summer	4.69
2.4	Ridge-Trough	Transition	0.66
3*			21.26
3.1	Trough-Ridge	Transition	11.19
3.2	Ridge-Trough	Winter	4.57
3.3	Ridge-Trough (high amplitude)	Transition	5.49
4	Zonal-Barotropic	Winter	14.23
5	Zonal	Summer	11.28
6	Trough-Ridge-Trough (northerly displacement)	Summer	6.59
7	Inverted Trough	Winter	1.82
8	Zonal-Baroclinic	Winter	0.53
9	Trough-Ridge-Trough (high amplitude)	Summer	0.31
10	Trough-Ridge-Trough (high amplitude)	Summer	0.05
11	Ridge-Trough (northerly displacement)	Winter	0.04

*subclusters are denoted by the parent cluster number followed by an additional digit identifying the subcluster.

Table 3. First and second most frequently preceding clusters for the 16 ULSI clusters show that the classification is generally temporally continuous.

Cluster	1 st Most Frequently Preceding Cluster		2 nd Most Frequently Preceding Cluster	
		(percent this cluster precedes)		(percent this cluster precedes)
1	1	(61.3%)	4	(25.3%)
2.1	2.1	(48.3%)	5	(13.5%)
2.2	2.2	(52.9%)	2.1	(10.9%)
2.3	2.3	(45.4%)	2.1	(21.6%)
2.4	2.4	(43.8%)	2.2	(20.8%)
3.1	3.1	(39.6%)	4	(20.2%)
3.2	3.2	(39.2%)	4	(20.2%)
3.3	3.3	(38.5%)	4	(16.3%)
4	4	(47.6%)	1	(17.8%)
5	5	(42.6%)	2.2	(17.2%)
6	6	(57.4%)	5	(30.8%)
7	7	(45.9%)	1	(44.7%)
8	8	(59.0%)	7	(23.1%)
9	6	(64.4%)	9	(26.7%)
10	6/9	(37.5%)	10	(25.0%)
11	8	(83.3%)	11	(16.7%)