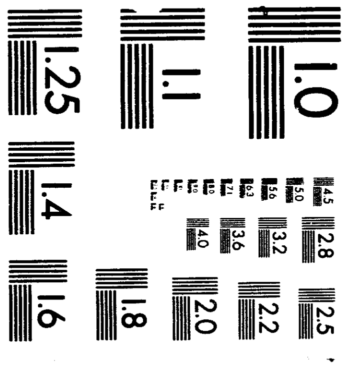


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3

1 of 1

Conf-931145--1

PNL-SA-20808

JAN - 4 1994
OSTI

SELECTING REPRESENTATIVE CLIMATE STATIONS
FOR USE IN A BUILDING ENERGY MODEL

D. L. Hadley

November 1993

Presented at the
CLIMA 2000 Conference
November 1-3, 1993
London, England

Work supported by
the U.S. Department of Energy
under Contract DE-AC06-76RLO 1830

MASTER

Pacific Northwest Laboratory
Richland, Washington 99352

SELECTING REPRESENTATIVE CLIMATE STATIONS FOR USE IN A BUILDING ENERGY MODEL

Donald L. Hadley
Senior Research Scientist
Pacific Northwest Laboratory
Richland, Washington 99352 U.S.A.

An energy impacts model is being refined to support ongoing development of major energy conservation standards for U.S. commercial buildings. When completed, the model will be used to evaluate potential impacts (energy savings and associated costs) of implementing the proposed standards. To work as intended, the model must contain a set of climate stations to represent the wide range of climatic conditions that occur across the United States.

Researchers developed a procedure that employs a user-selectable climate database 1) to objectively identify, using a clustering technique, a unique set of climate zones for a specified geographical area, and 2) to specify the single most representative station for each climate zone. The process provides a more objective, technically sound basis for selecting climate zones and stations, thereby minimizing researcher bias. The procedure and its application to U.S. energy conservation standards development activities are described in this paper.

La Sélection des Stations Climatiques Representatives en Construisant une Representation de l'Utilisation d'Énergie

Afin d'établir des étalons de conservation d'énergie pour des bâtiments commerciaux aux États Unis, des savants utilisent une representation pour rechercher leurs effets sur la consommation d'énergie. Cette representation

doit comprendre des stations climatiques qui representent la grande diversitée des conditions climatiques aux Etats Unis.

Des savants developerent une méthode par laquelle on pourrait organizer l'information au sujet d'un climat particulier. En utilisant cette information, on pourrait 1) identifier les zones climatiques uniques pour une region geographique, et 2) preciser la station la plus representative de chaque zone climatique. En utilisant cette methode, des recherches climatiques deviendront plus objectives et scientifiquement bien fondées. Cette méthode, et son application au dévelopment des étalons de conservation d'énergie, sont les sujets de ce bulletin.

Auswahl Repräsentativer Klimastationen zum Gebrauch in einem Energieeinwirkungsmodell fur Gebäude

Ein Einwirkungsmodell der Energie wird raffiniert, um die laufende Entwicklung bedeutender Energiesparungsnormen für Handelsgebäude in den Vereinigten Staaten zu unterstützen. Wenn das Modell fertiggestellt ist, wird es gebraucht, um mögliche Einwirkungen (Energiesparung und damit verbundene Kosten) bei der Ausführung der vorgeschlagenen Normen auszuwerten. Soll das Modell wie vorgesehen arbeiten, muss es eine Reihe von Klimastationen einschliessen, damit der grosse Bereich von klimatischen Bedingungen, die in den Vereinigten Staaten vorkommen, dargestellt ist.

Forscher entwickelten ein Verfahren, mit dem der Verbraucher die Klimadatenbasis auswählt, um 1) objektiv, unter Benutzung einer Gruppierungstechnik, eine kennzeichnende Reihe von Klimazonen für ein bestimmtes geographisches Gebiet zu identifizieren, und 2) die repräsentativste Station für jede Klimazone zu spezifizieren. Das Verfahren bietet eine objektivere, technisch einwandfreie Basis, um Klimazonen und Stationen zu wählen, und damit Vorurteile der Forscher zu reduzieren. Das Verfahren und dessen Anwendung für die Entwicklung der Energiesparungsnormen in den Vereinigten Staaten werden hier beschrieben.

1.0 INTRODUCTION

The U.S. Department of Energy (DOE) and the American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. (ASHRAE) are in a continual cycle of revising and updating their respective commercial building energy standards. These standards [1,2] specify energy-efficiency measures (EEMs) that are technologically feasible and economically justifiable (i.e., that are cost-effective over a building's life cycle). To help DOE and ASHRAE assess the impacts on energy consumption and energy costs, the Pacific Northwest Laboratory is refining a model that estimates changes in nationally aggregated building energy-use intensities (EUIs) resulting from proposed changes to the energy standards. One prerequisite for this model is the identification of a set of climate zones and representative climate stations covering the full range of unique climatic conditions occurring over the contiguous United States.

The impact model is based on parametric computer analyses of the hourly simulation of prototypical building performance for one or more climates. Regional EUIs generated from the model are then aggregated, with appropriate weighting for population, building growth, and other factors, to a single national EUI. For the results of the model to be applicable at the national level, however, the initial choice of climate characteristics for each region must cover all of the representative climates.

Unfortunately, there is no consensus as to the correct number of climates or the delineation of climate zones. The identification of the specific climate zones and/or the selection of representative climate stations are frequently without solid technical justification and arbitrary, based on the researcher's previous experience, and, all too often, limited by availability of data. For example, a recent evaluation [3] of the effectiveness of the DOE commercial building energy standard was based on 10 commercial building types in six U.S. locations (Seattle, Washington; Los Angeles, California; El Paso, Texas; Lake Charles, Louisiana; Washington, D.C.; and Madison, Wisconsin). These six were a subset of the eight stations originally identified by Bowen et al. [4] as encompassing the full range of climate variations found in the contiguous United States. However, other researchers have used different stations as being "representative."

Andersson et al. [5] described a methodology by which relatively homogeneous climate regions appropriate to a specific application -- in this case, building energy -- could be generated. The basic premise of the work is that it is possible to objectively aggregate specific U.S. population centers into common groups based on climate elements that most directly influence building energy consumption. The three elements selected were temperature, humidity, and solar radiation, represented as four variables -- heating degree-days (base 65°F), cooling degree-days (base 65°F), latent enthalpy hours (base 61°F dew-point temperature and 75°F dry-bulb temperature), and an annual average clearness index. The authors presented the results of a number of aggregations ranging from 5 to 24 regions. Clearly, 5 regions were too few, as many dissimilar sites were forced into the same climate region designation. A more reasonable number of climate regions appeared to be in the range of 11 to 15. Twenty-four climate centers produced a number of single-station climate regions. A limitation of this methodology is that the user selects an initial set of climate centers around which all other stations are aggregated.

The specific sites and the number of climate centers are arbitrary and user-specified. Effective user specification requires a *priori* knowledge of relevant climate regions.

This paper presents a procedure referred to as synoptic climatology indexing for 1) objectively defining a unique set of climate zones for a specified geographical area, and 2) identifying the single most representative climate station for each zone, with the underlying climate database defined by the user. The procedure provides a sound technical basis for the selection process and does not require previous knowledge of the climate classification.

2.0 METHODOLOGY

The synoptic climatology indexing procedure uses a hierarchical clustering technique to group climate stations into climatologically homogeneous groups. A group, or cluster, of stations is considered homogeneous if its members are more similar, on average, to others in the cluster than to those in other clusters. This methodology is similar to that used by Andersson et al. [5] to aggregate U.S. population centers, except that it is not necessary to initially specify a climate center around which other sites are aggregated.

This methodology is based on the objective synoptic climatology classification methodology first introduced by Kalkstein and Corrigan [6] and since used widely in a number of diverse applications. Hadley [7] applied the procedure to identify homogeneous weather day-types in an analysis of daily variation in heating and air-conditioning energy use. Kalkstein et al. [8] used the approach to show signs of climatic change in the western North American Arctic.

Synoptic Climatology Indexing

This methodology uses a hierarchical clustering technique with a distance metric computed using an "average linkage." This technique compares the average squared Euclidean distance [9] between all possible pairs of observations between all combinations of stations/clusters. The clustering process is iterative; two stations/clusters with the minimum average distance are merged to form a single new cluster at each iteration. The merging continues until the final two are combined in the final iteration. The average linkage computation is designed to minimize within-cluster variance and maximize between-cluster variance.

A number of techniques are available to rigorously determine the number of clusters naturally occurring in the data set [6,9]. In theory, there is an identifiable breakpoint at which dissimilar (unlike) clusters begin to be merged and the within-cluster variance increases sharply. The "correct" number of natural clusters occurs just prior to this point. If more than the "correct" number of clusters are generated, the natural clusters will be artificially split into two or more similar clusters. If the number of clusters decreases below the "correct" number, two or more dissimilar clusters are forced together. In practice, this point is not easily identified. Consequently, the final number of climate zones may be selected on the basis of other factors (e.g., population density, construction activity).

The clusters contain stations that are climatologically similar and that constitute a unique climate zone.

The final step in the selection process is a ranking of all stations in a climate zone to determine the single station that best represents the zone's mean climatic characteristics. The metric used for this ranking is a composite Z-score statistic. A Z-score of 0.0 means that the variable for that station is equal to the average of that variable; a Z-score of 1.0 means that it is one standard deviation from the mean.

The composite Z-score (Z_c) is the average of the absolute value of the individual Z-scores (Z_i) for that station. It is computed as follows:

$$Z_c = \Sigma \mid Z_i \mid / n$$

where $Z_i = (X_i - u) / \sigma$

X_i = individual climate variable

u = mean

σ = standard deviation

n = number of variables.

The station with the lowest composite Z-score is the station most representative of the climate for that region.

Climate Characterization

Typical Meteorological Year (TMY) data for 209 continental U.S. stations were used in this analysis. Because of the restriction in the impact model to develop climate zone designations consistent with the regional building characteristics data available only for the four major U.S. Census Bureau regions (West, North Central, Northeast, and South), each TMY station was first assigned to the appropriate region.

The climate of each station was characterized by those variables that were used to develop the Alternative Component Package tables in the DOE and ASHRAE commercial building standards [1,2]. The 11 variables comprising the climate database were

1. heating degree-days, base 18.3°C (65°F, HDD65)
2. heating degree-days, base 10.0°C (50°F, HDD50)
3. cooling degree-days, base 10.0°C (50°F, CDD50)
4. cooling degree-days, base 18.3°C (65°F, CDD65)
5. cooling degree-hours, base 26.7°C (80°F, CDH80)
6. daily incident solar radiation on a north vertical face (VSN), W/m^2 ($Btu/ft^2/day$)
7. daily incident solar radiation on an east or west vertical face (VSEW), W/m^2 ($Btu/ft^2/day$)
8. daily incident solar radiation on a south vertical face (VSS), W/m^2 ($Btu/ft^2/day$)
9. diurnal range of temperature in the warmest month (DR), °C (°F)
10. number of hours between 8:00 a.m. and 4:00 p.m. with temperature <12.8°C (<55°F, T55)
11. number of hours between 8:00 a.m. and 4:00 p.m. with

temperature between 12.8°C and 20.5°C (55°F and 69°F, T69).

In this analysis, all variables were equally weighted.

3.0 RESULTS

The synoptic climatology indexing technique was applied to the selected climate database. This analysis was performed initially for the continental U.S. stations only, so stations in Alaska, Hawaii, and the islands of the Caribbean and South Pacific were not included.[1] In our situation, final selection of the number of climate zones used subjective judgment, based on an analysis of different levels of clustering (ranging from 5 to 30 zones) and an examination of the average distance between clusters merged at each successive iteration.

We were constrained by the need to limit the climate zones to a manageable number. Subsequent work in finalizing the national energy model would require parametric simulations of hourly building energy use for up to 10 different building types for each climate zone and for a variety of building parameters. Obviously, the fewer the climate stations, the fewer the simulations needed to complete the model.

Resultant Climate Zones

[1] Addition of these stations outside the continental U.S. to the cluster analysis is not expected to significantly affect the resulting clusters.

The synoptic climatology indexing resulted in the identification of a unique set of climate zones for each of the four census regions. The final number of different climate zones selected ranged from two in the Northeast region to six in the West region. The number of climate zones is dependent on the climatic diversity of the region. The extent of each zone is shown in Figure 1. To delineate the approximate geographic extent of each of the climate zones, a line has been loosely drawn between adjacent stations in different climate zones. However, extreme care must be taken in the extrapolation of climate zones in mountainous terrain.

Climate Station Ranking and Selection

A climatologically representative station for each of the climate zones was determined using a composite Z-score, identifying the one station whose climate characteristics most closely fit the average of all stations in that climate zone. To illustrate the process, climate parameters and individual Z-Scores are listed in Tables 1 and 2 for 15 stations in the South region, Climate Zone 8. These stations have been ranked is ascending order of composite Z-score.

Because the intent of the energy model is to estimate national impacts of proposed changes to EEMs in commercial buildings, the final climate station selection also included a consideration of population and new building construction. Stations in remote, low-population areas with little new construction activity were not included in the final station selection. For example, in the West census region, Denver was selected over Cedar City

(Climate Zone 11), and Los Angeles was selected over both Sunnyvale and Santa Maria (Climate Zone 15).

4.0 CONCLUSION

The synoptic climatology indexing procedure described in this paper is successful in objectively identifying unique climate zones. The procedure eliminates many of the limitations and personal biases introduced by the more common subjective approaches to identifying climate zones and selecting representative climate stations. Although developed specifically for application to data for the continental United States, the methodology has global applicability whenever it is necessary to define a set of climate zones. The technique is "data-neutral"; that is, it is viable regardless of the underlying climate variables selected or whether the database is regional, national, or global. The user is free to select climate variables appropriate for the intended application, and can apply weights to the individual variables to emphasize a specific climate element or elements, depending on the goals of the application.

REFERENCES

[1] U.S. Department of Energy. 1989. *10 CFR 435 Subpart A, Energy Conservation Voluntary Performance Standards for Commercial and Multi-Family High Rise Residential Buildings; Mandatory for New Federal Buildings; Interim Rule*, 54 FR 4554, January 30, 1989. Superintendent of Documents, U.S. Government Printing Office, Washington, D.C.

[2] American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. 1989. *ASHRAE/IES Standard 90.1-1989*, "Energy Efficient Design of New Buildings Except Low-Rise Residential Buildings." ASHRAE, Atlanta, Georgia.

[3] Hadley, D.L., and M.A. Halverson. 1992. *Energy Conservation Potential of the U.S. Department of Energy Interim Commercial Building Standards.* PNL-7967, Pacific Northwest Laboratory, Richland, Washington.

[4] Bowen, W.M., S. Moreno, and A.R. Olsen. 1982. *Climate Selection and Development of Climate Indicators.* PNL-4435, Pacific Northwest Laboratory, Richland, Washington.

[5] Andersson, B., W.L. Carroll, and M.R. Martin. 1986. "Aggregation of U.S. Population Centers Using Climate Parameters Related to Building Energy Use." *J. Climate Appl. Meteor.* 25:596-614.

[6] Kalkstein, L.S., and P. Corrigan. 1986. "A Synoptic Climatological Approach for Geographical Analysis: Assessment of Sulfur Dioxide Concentrations." *Ann. Assoc. Amer. Geogr.* 76:381-395.

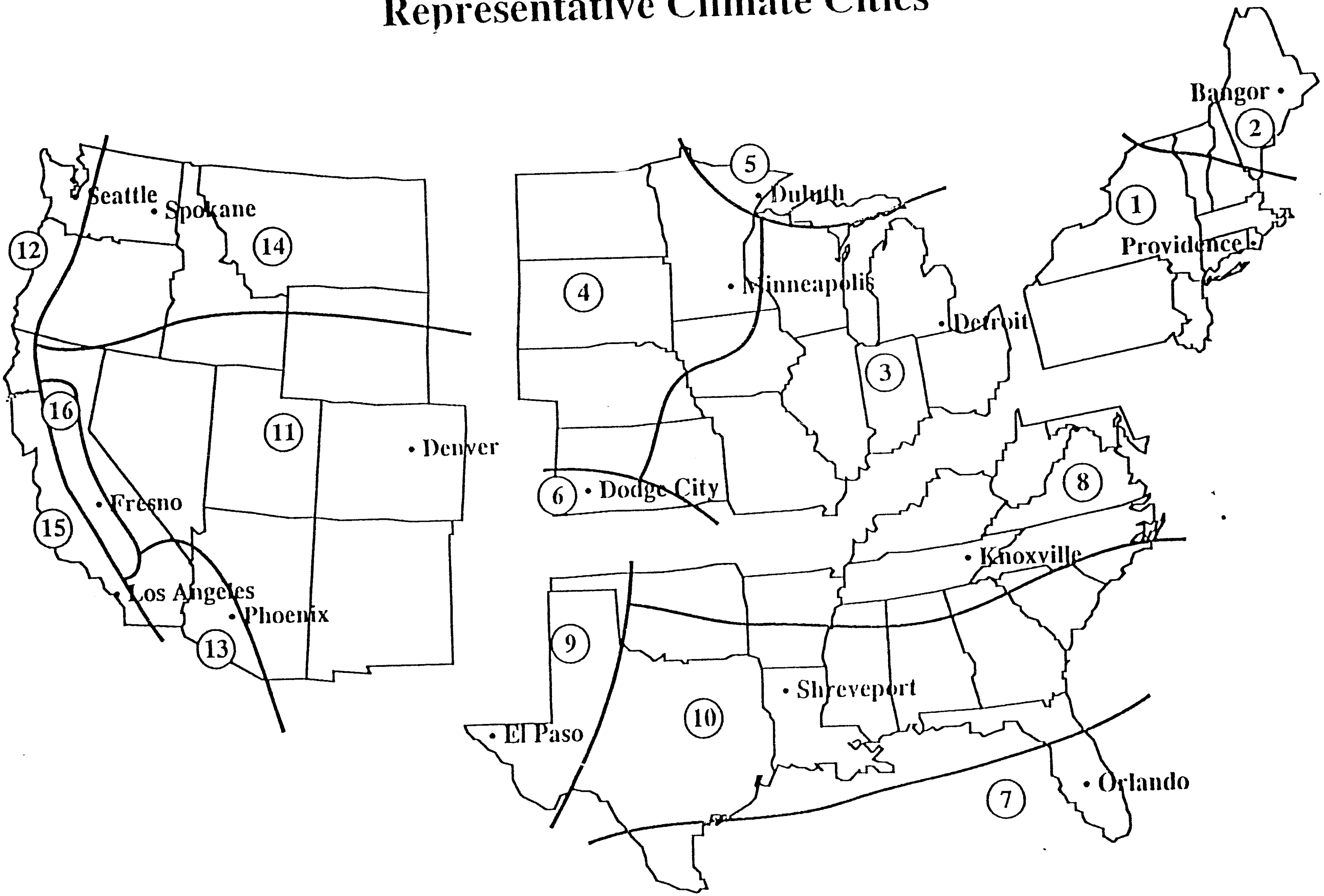
[7] Hadley, D.L. 1992. "Daily Variations in HVAC System Electrical Energy Consumption in Response to Different Weather Conditions." *Energy and Buildings* IN PRESS.

[8] Kalkstein, L.S., P.C. Dunne, and R.S. Vose. 1990. "Detection of Climatic Change in the Western North American Arctic Using a Synoptic Climatological Approach." *Journ. of Climate* 3:1153-1167.

[9] Kalkstein, L.S., G. Tan, and J.A. Skindlov. 1987. "An Evaluation of Three Clustering Procedures for Use in Synoptic Climatological Classification." *J. Climate Appl. Meteor.* 26:717-730.

Figure 1.

Representative Climate Cities



ESD9207001.07

Table 1. Climate variables for the top 15 cities in Climate Zone 8, South Census Region

CITY	Climate Parameter										
	HDD50	HDD65	VSN	VSEW	VSS	CDD50	CDD65	CDH80	DR	T55	T69
Knoxville, TN *	713	2121	59	100	118	2475	841	2133	9.9	1076	703
Richmond, VA	734	2164	57	98	121	2347	735	2234	9.8	996	716
Raleigh, NC *	628	1949	58	102	123	2492	772	2054	9.2	918	740
Norfolk, VA *	658	2005	58	104	127	2576	881	2530	8.3	1014	685
Greensboro, NC	701	2089	59	106	131	2374	721	2023	9.7	1018	718
Chattanooga, TN *	684	1997	58	97	114	2584	856	2822	9.8	1050	684
Nashville, TN *	647	2005	58	98	113	2546	862	2821	10.1	897	749
Roanoke, VA	844	2329	57	100	124	2214	657	1837	10.6	1148	713
Patuxent, MD	788	2223	56	100	124	2322	716	1648	7.2	1118	729
Charlotte, NC *	603	1896	60	106	127	2610	861	2388	10.9	892	777
Atlanta, GA *	481	1706	61	106	122	2687	870	2111	9.8	915	749
Louisville, KY *	1028	2522	56	96	116	2302	754	2620	9.8	1192	636
Greenville, SC *	504	1789	60	107	128	2535	778	1941	9.8	866	851
Washington, DC *	1113	2682	55	95	119	2074	602	1996	10.3	1265	657
Asheville, NC	782	2335	59	103	124	1912	424	721	11.7	1083	915
Average all sites	758	2143	58	102	123	2466	809	2535	9.[illegible]	1028	709
Std. Deviation	245	386	2	7	7	330	198	1342	1.4	164	87

* Indicates a major metropolitan area with total population exceeding 250,000.

Table 2. Individual and composite Z-scores for the top 15 ranked cities in Climate Zone 8, South Census Region.

City	Climate Parameter											Composite Z-score
	HDD50	HDD65	VSN	VSEW	VSS	CDD50	CDD65	CDH80	DR	T55	T69	
Knoxville, TN *	-0.18	-0.06	0.21	-0.26	-0.65	0.03	0.16	-0.30	0.02	0.29	-0.07	0.20
Richmond, VA	-0.10	0.05	-0.65	-0.60	-0.21	-0.36	-0.37	-0.22	-0.06	-0.20	0.08	0.26
Raleigh, NC *	-0.53	-0.50	0.16	-0.02	0.01	0.08	-0.19	-0.36	-0.49	-0.67	0.36	0.31
Norfolk, VA *	-0.41	-0.36	0.05	0.35	0.52	0.33	0.36	-0.00	-1.08	-0.09	-0.28	0.35
Greensboro, NC	-0.23	-0.14	0.37	0.71	1.05	-0.28	-0.45	-0.38	-0.14	-0.06	0.10	0.36
Chattanooga, TN *	-0.30	-0.38	0.10	-0.74	-1.17	0.36	0.24	0.21	-0.06	0.13	-0.29	0.36
Nashville, TN *	-0.45	-0.36	0.05	-0.52	-1.27	0.24	0.27	0.21	0.17	-0.80	0.46	0.44
Roanoke, VA	0.35	0.48	-0.49	-0.24	0.20	-0.76	-0.77	-0.52	0.49	0.73	0.05	0.46
Patuxent, MD	0.12	0.21	-0.71	-0.34	0.15	-0.43	-0.47	-0.66	-1.91	0.55	0.23	0.53
Charlotte, NC *	-0.63	-0.64	0.75	0.69	0.59	0.44	0.26	-0.11	0.72	-0.83	0.78	0.59
Atlanta, GA *	-1.13	-1.13	1.34	0.65	-0.08	0.67	0.31	-0.32	-0.06	-0.69	0.46	0.62
Louisville, KY *	1.10	0.98	-0.98	-0.96	-0.92	-0.49	-0.28	0.06	-0.06	1.00	-0.84	0.70
Greenville, SC *	-1.04	-0.92	0.91	0.79	0.65	0.21	-0.16	-0.44	-0.02	-0.99	1.64	0.71
Washington, DC *	1.45	1.40	-1.25	-1.02	-0.53	-1.18	-1.05	-0.40	0.33	1.08	-0.60	0.93
Asheville, NC	0.10	0.50	0.37	0.14	0.20	-1.68	-1.94	-1.35	1.31	0.33	2.37	0.94

* Indicates a major metropolitan area with total population exceeding 250,000.

DATE
FILMED
2/17/94
END