JP1.11 ANNUAL COURSE OF SUCCESSIVE 30-DAYS' OVERALL, ABOVE NORMAL, AND BELOW NORMAL TEMPERATURE PERSISTENCE AT ONE-DAY INTERVALS FOR FOUR U.S. STATIONS WITH LENGTHY HISTORIES

By Charles J. Fisk*
NAWCWPNS- Pt. Mugu, CA

1. INTRODUCTION

The Pearson correlation coefficient is frequently used for measuring linear temperature association (persistence) between adjacent time-series. While the percent of variance explained might be meager and the predictive value poor, the correlation magnitude still serves as an objective, comparative means of characterizing different calendar-periods' (e.g., January vs. February, February vs. March) persistence. Unlike frequency persistence (percent incidence of successive periods with the same anomaly sign), the correlation is sensitive to anomaly departure magnitudes, especially extreme ones.

When the data are in standardized form, the correlation is simply the sum of the products of the individual z-scores divided by the sample size (n). By substituting z-score product summations from appropriate subsets in the numerator while retaining n in the denominator, the original coefficient can also be mechanically decomposed into "above normal" and "below normal" persistence magnitudes. Above (below) normal persistence would be the sum of the z-score products for cases in which the leading time-series had a positive (negative) z-score; each of these totals would be divided by n. The sum of the "above normal" (or "warm") and "below normal" (or "cold") persistence values would yield the original correlation coefficient. A large disparity in magnitudes between the two might suggest of a natural inherent tendency of one type predominating over the other.

The retention of (total) n in the denominator for above and below normal persistence magnitude calculation assumes that leading time-series subset frequencies are approximately equal (a likely fact if they come from a normally distributed parent series), and that actual discrepancies, especially with regards to long periods of record, will have negligible effects on results (e.g., warm/cold persistence magnitudes and the arithmetic differences between them) interpretation. If subset frequencies are identical, n/2 can be substituted into the denominator and the results further divided by 2 to yield the same component figures as if n was used alone.

* *Corresponding author address:* Charles J. Fisk, NAWCWPNS, Point Mugu, CA. 93042: e-mail: fiskcj@navair.navy.mil

2. PURPOSE

Using this approach, the annual course of successive 30-days' temperature persistence (overall, above normal, and below normal) is explored at one-day intervals (moving array of 365 correlations – Feb. 29 is averaged with Feb. 28 for leap years) for four U.S. stations with extended periods of record. These are New York City Central Park (125 years), Minneapolis-St. Paul (128 years), Salt Lake City (71 years), and Los Angeles Civic Center (80 years). The analysis is done with linear trend removed. To gauge the rough significance of the above normal/below normal persistence magnitude discrepancies, a table of standard errors is created and utilized based on analysis of simulated data sets with varying built-in correlations and overall sample sizes (n).

3. ABOVE NORMAL VS. BELOW NORMAL PERSISTENCE CONTRASTS

Generation of a means to evaluate the statistical significance of above normal persistence/below normal persistence magnitude contrasts was made possible by 1) the availability of software that can generate user-defined covariance matrices and 2) the statistical principle that defines correlations as being equal to covariances in a standardized data set. A bivariate data set of a chosen correlation ρ and sample size n could thus be easily generated by assuming that the "covariance" was really a "correlation." As such, the resulting x and y "observations" would each be z-scores, and summation products could be manipulated into the "above normal"/ "below normal" difference statistics described above. Analysis of a large number of simulated data sets of this kind with predetermined sample sizes and correlations should permit construction of a standard error type table that could yield approximate t-value statistics for the differences. To this end, a series of data sets with built-in correlations of zero, +0.25, +0.50, and +0.75; and 5000-group sample sizes of 25, 50, 75, 100, and 150 were generated using the FACTOR module of the SYSTAT statistical software package. For example, for the n=150 and ρ=+.250 combination, 750,000 original cases with an overall or "universe" correlation of +.25 were created. Then, the set was subdivided into 5000 groups of 150 members each. Using another module, within-group standardizations were performed; above normal, below normal, and arithmetic differences then calculated using the subset product summation method described above. Standard deviation of the 5000 arithmetic differences was next determined, representing a standard error figure for the given n and ρ values in question.

In the case of n=150 and ρ=+.250, the absolute range of differences was -.177 to +.180, the standard deviation of these differences 0.050. This figure would thus serve as the standard error for the n=150 and ρ=+.250 combination, as Table 1 below indicates. Inspecting the Table further, the standard error magnitudes are dependent much more strongly on n (figures at n=150 less than half those at n=25) than ρ (figures at ρ=+.750 about 5% higher than those at ρ=+.250), so standard error figures selected or interpolated on the basis of sample r's, not universe ρ's, could be applied on observed persistence component differences with only slight loss of certainty regarding true standard errors (and t-values).

In addition, the simulated observational groupings of z-scores and z-score products that were used to derive Table 1's standard error statistics were composed, like real-world data sets, of mostly unequal lead series subset frequencies (always unequal when n was odd-numbered), so this variability was incorporated into the standard error figures.

Table 1: Standard Errors for Absolute Differences Between Above Normal and Below Normal Persistence Magnitudes, by Sample Size (n) and Universe Correlation (ρ) – Simulated data

	ρ=.00	ρ=.25	ρ=.50	ρ=.75
n=25	.115	.115	.117	.120
n=50	.084	.084	.087	.088
n=75	.069	.070	.071	.072
n=100	.060	.060	.061	.063
n=150	.049	.050	.051	.052

Persistence maxima would likely identify times of the year in which temperatures are particularly affected by combinations of favorable circulation patterns enhancing the effects of abnormally forward or backward surface (ground or ocean) conditions locally or in air-mass source regions, or generally more sluggish patterns of summer in which abnormal thermal patterns are more likely to stay entrenched. The relative minima likely reflected preferred seasonal break times in temperature between the winter-to-summer and summer-to-winter transitions.

As these are moving arrays of correlations, the same caution in interpreting "peaks" and "valleys" applies as for ordinary moving averages, especially with regards to those that appear in overlapping sequences.

4. SIXTY-DAY TEMPERATURE PERSISTENCE FOR NEW YORK CENTRAL PARK

Figures 1 to 3 trace the day-to-day statistics of successive 30 days' Overall Persistence, Above (Below) Normal Persistence, and Above/Below Normal Persistence differences, respectively, for New York Central Park, based on data from the start of 1876 through February 2001. For each graph, the 365 individual points are centered on the first day of the second (lagged) 30-day period. For instance, a statistic positioned on July 25 would indicate the magnitude for the June 25 to July 24 vs. July 25 to August 23 sequence, the sequence visualized by extending imaginary brackets about one-month forward and backward from that point. For descriptive convenience, the sequence could also be described as "centered" on July 25, or having a "midpoint" on that date.

Average overall persistence for Central Park is r=+.213 (horizontal line in Figure 1), the lowest of the four stations in this report. According to the NCDC International Station Meteorological Climatic Summary, the station "is close to the path of most storm and frontal passage systems which move across the North American Continent". As such, "frequent passage of weather systems often helps reduce the length of warm and cold spells". This likely applies both to short-range and long-range time scales and probably explains the low average correlation.

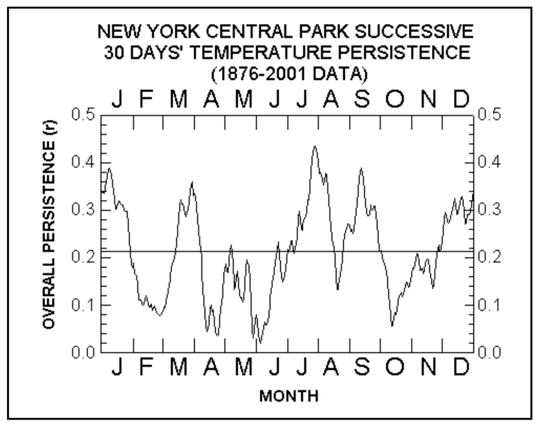

Figure 1. Successive 30 Days' Overall Temperature Persistence Magnitudes for New York Central Park, New York

Still, the calendar day to calendar day pattern is far from uniform. Four relative maxima are evident. The first (r=+.389 for December 9 to January 7 vs. January 8 to February 6) describes a mid-winter relative maximum, peak two (r=+.359 for February 28 to March 29 vs. March 30 to April 28) an early spring one. The third peak (r=+.436 for June 29 to July 28 vs. July 29 to August 27 (the absolute maximum for the year) is the summer maximum, likely reflecting the positional influences of the Azores-Bermuda High in promoting long-term anomalous temperature regimes. Peak four (r=+.388 for August 13 to September 11 vs. September 12 to October 11), covering late-summer to early fall, has its time frame overlapping five days with that of peak three. Examining Figure 2, which shows the breakdown into above normal and below normal components, the early spring and late summer/early autumn peaks are mostly above normal in makeup (dotted line noticeably higher than the solid line for the sequences in question), the mid-winter and mid-summer sequences showing roughly equivalent magnitude decompositions. Interpreted, the early spring warm persistence relative proclivity might be an effect of forward (premature departure of snow) ground conditions in air-mass source regions, or possibly even abnormal far away oceanic influences (El Nino?), producing the potential for more "intense" 60-day anomalies in mean temperature when long-term circulation patterns are favorable. The late-summer/early-fall discrepancy might relate to a seasonal-lag effect. Inspecting Figure 3, the discrepancy magnitudes are around +0.10. Taking into account the 125-year period of record, overall r's, and interpolated standard errors from Table 1 (~0.55), this translates into t-values of about 1.8, significant for single comparison one-tail tests at the .05 level, but probably not conclusive given the large number of total correlations under consideration in the graph (multiple comparisons issue).

Four relative minima are also present. In order of highest, the first (r=+.078 for January 27 to February 25 vs. February 26 to March 27) might reflect a comparatively favored onset time for the winter to spring transition (late February break-point), the second (r=+.057 for September 22 vs. October 11 vs. October 12 to November 10) the same for summer to fall (mid-October breakpoint). Part of a lengthy group of contiguous spring sequences that display the sixteen lowest r's of the year, the two others include r=+.036 for March 25 to April 23 vs. April 24 to May 23, and r=+.021 for May 6 to June 4 vs. June 5 to July 4, the latter the extreme minimum for the year. For the March 9 to April 7 vs. April 8 to May 7 through June 1 to June 30 vs. July 1 to July 30 sequences -- a run of 85, persistence is below average for all but four.

From Figure 2, the spring sequences' overall lack of temperature persistence would be even more striking except for the fact that there is a relatively pronounced cold persistence component for those with midpoints in early to mid-May. Highest cold persistence magnitude over this interval is for the April 8 to May 7 vs. May 8 to June 6 sequence (+.178), the corresponding warm persistence statistic just (+.034). Figure 3 indicates that a number of magnitude contrasts, including this one, are much more negative than -0.10. That for the April 10 to May 9 vs. May 10 to June 8 sequence (-.169), the most extreme, has an estimated t-value of –3.1, significant at approximately the .001 level for a one-tail test. From this, perhaps there really is an inherent tendency for below normal 60-day persistence at New York Central Park over the early April to early June frame to be more "intense" than above normal. For many of the other individual sequences that include the run of 85, however, the above normal and below normal figures are each around zero.

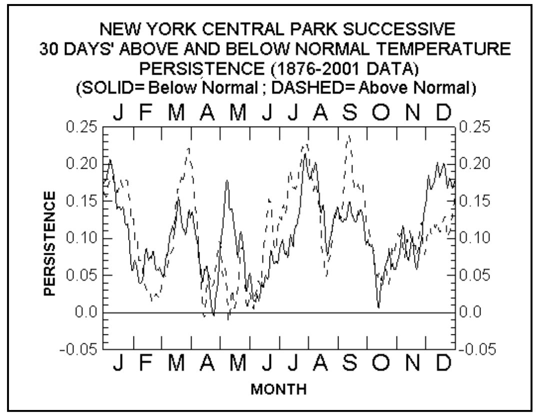

Figure 2. Successive 30 Days' Above and Below Normal Temperature Persistence Magnitudes for New York Central Park, New York

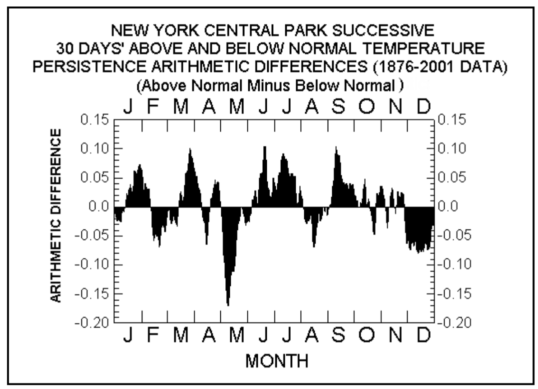

Figure 3. Successive 30 Days' Above and Below Normal Temperature Persistence Magnitude Differences for New York Central Park, New York

5. SIXTY-DAY TEMPERATURE PERSISTENCE FOR MINNEAPOLIS-ST. PAUL

Figures 4 to 6 trace the day-to-day statistics of successive 30 days' Overall Persistence, Above (Below) Normal Persistence, and Above/Below Normal Persistence differences, respectively, for Minneapolis-St. Paul, Minnesota, based on data from the start of 1873 through February 2001. The NCDC narrative describes the climate as "predominantly continental" with "large temperature variations". Minneapolis-St. Paul also typically has a snowcover from late November to early March, the abnormal absence or premature/prolonged presence of which can influence temperature persistence.

From Figure 4, Minneapolis-St. Paul's average overall persistence (r=+.241, indicated by the horizontal line) is somewhat higher than Central Park's. Winter to early-spring overall magnitudes are above average for 89 consecutive sequences, from the January 4 midpoint to April 2's.

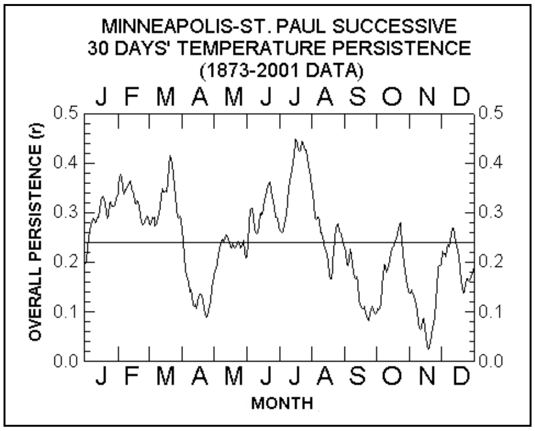

Figure 4. Successive 30 Days' Overall Temperature Persistence Magnitudes for Minneapolis-St. Paul, Minnesota

Among these are relative maxima for the slightly overlapping sequences January 4 to February 2 vs. February 3 to March 4 (r=+.377) and February 19 to March 20 vs. March 21 to April 19 (r=+.415). Most prominent peak of the year, like that for Central Park, is in mid-summer, the absolute maximum (r=+.449) noted for the June 17 to July 16 vs. July 17 to August 15 sequence. Starting with the June 3 centered sequence, overall persistence is higher than average for 70 consecutive sequences through August 11. Fall to early winter persistence is relatively low, more than 90% of the sequences exhibiting figures below +.241; the relative peaks are barely above this overall mean.

Three relative minima are present. The first (r=+.090 for the March 25 to April 23 vs. April 24 to May 23 sequence is suggestive of a spring transition breakpoint, the second (r=+.083 for the August 25 to September 23 vs. September 24 to October 23) a summer to fall one. The third minimum, +.026 for October 19 to November 17 vs. November 18 to December 17 (absolute lowest) might be a fall to winter "divide".

Figure 5 shows the breakdown into above normal and below normal components. Except for eighteen contiguous late January to early February midpoints (mid-winter), above normal magnitudes exceed below normal ones for each of the 110 sequences from the December 18 through April 6 midpoints, inclusive. Highest warm persistence statistic is for February 19 to March 20 vs. March 21 to April 19 (+.253); this is also the absolute maximum for the year. Cold persistence magnitude for January 4 to February 2 vs. February 3 to March 4 (+.230) is also the extreme yearly maximum for this type. The sequences centered in mid-to-late July also show high relative warm and cold persistence magnitudes each.

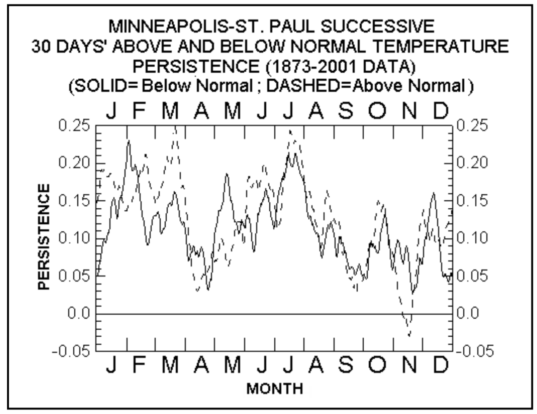

Figure 5. Successive 30 Days' Above and Below Normal Temperature Persistence Magnitudes for Minneapolis-St. Paul. Minnesota

Inspecting Figure 6, a number of the winter and early-spring sequences' warm persistence vs. cold persistence excesses approach or even pass +0.10. Again, while significant at about the .05 level for a one-tail test, the differences are probably not conclusive statistically given the large number of total points in the graph. However, allowing for this fact, the relative inclination might be an effect of unusually snow-deficient or forward surface conditions locally and in air-mass source regions. The effect of strong El Ninos in producing occasionally much milder than normal temperatures in the northern tier of Midwestern states, especially in February and March, is also well known.

Interestingly, the most pronounced above normal below normal contrast is situated just five sequences later than Central Park's, and is of the same sign (negative). For the April 15 to May 14 vs. May 15 vs. June 13 sequence at Minneapolis-St. Paul, above

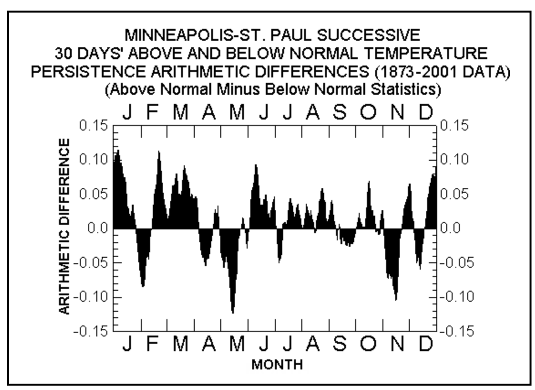

Figure 6. Successive 30 Days' Above and Below Normal Temperature Persistence Magnitude Differences for Minneapolis-St. Paul, Minnesota

normal persistence (+.062) is appreciably lower than the corresponding below normal figure (+.185). The discrepancy (-.123) produces an estimated –2.2 t-value, significant at the .015 level for a one-tail test.

6. SIXTY-DAY TEMPERATURE PERSISTENCE FOR SALT LAKE CITY

Figures 7 to 9 trace the respective Overall Persistence, Above (Below) Normal Persistence, and Above/Below Normal Persistence difference statistics for Salt Lake City, Utah based on data from the start of 1930 through February 2001. The period of record is about 45% shorter than those of Central Park and Minneapolis-St. Paul. According to the NCDC International Station Meteorological Climatic Summary, the station "is located in a north Utah valley [at the eastern edge of the Great Basin] surrounded by mountains on three sides and the Great Salt Lake to the Northwest ". The mountains "to the north and east act as a barrier to frequent invasions of cold continental air" and the lake, "which never freezes over … can moderate cold winter winds from the northwest" .

From Figure 7, Salt Lake City's average overall persistence is r=+.250 (upper horizontal line), about the same as Minneapolis-St. Paul's. Three relative maxima are present, a clear-cut dominant one in winter, and two lesser ones, portions of spring and summer, respectively. Absolute maximum (r=+.552) appears for the December 14 to January 12 vs. January 13 to February 11 sequence, part of a run of 102 consecutive ones, from the November 27 to March 8 midpoints, that are above average. The other two peaks are r=+.374 for March 20 to April 18 vs. April 19 to May 18 and r=+.402 for June 6 to July 5 vs. July 6 to August 4. Beginning with the August 27 midpoint, persistence is below average for 87 consecutive sequences, through November 21. On an overall basis, the fall/early winter sequences clearly have the lowest overall persistence.

Four relative minima are evident. The first three, having only modestly depressed correlations, are: r=+.110 (centered on both March 22 and 23), r=+.130 (centered on June 1), and r=+.115 (centered on July 31). The fourth, r=-.029 for October 8 to November 6 vs. November 7 to December 6, is the lowest of the year and also the most negative encountered among the four stations. As the graph shows, however, this minimum is less a localized breakpoint than part of a lengthy series of generally low magnitudes that comprise Fall.

Figure 8 shows the breakdown into above normal and below normal components. Except for a few sequences centered in early January, the winter overall maximum is attributable mostly to exceptionally high magnitudes of cold persistence. For example, for the 33 consecutive sequences from December 26 to January 27 midpoints, cold persistence magnitudes are +.250 or higher for all but two. Absolute highest magnitude (+.302) is noted for the January 12 and 13 sequence midpoints. None of the other non-winter sequences are as high as +.200 for cold persistence. The +.250 warm persistence magnitude for the sequence midpoints January 9, 11, and 12, is also a yearly absolute maximum for this type; a secondary peak (+.242) is also noted for the July 2 midpoint. Perhaps the high winter cold persistence statistics relate to Salt Lake City's basin setting, and a tendency for outbreaks of cold, dense air to become topographically trapped for extended periods.

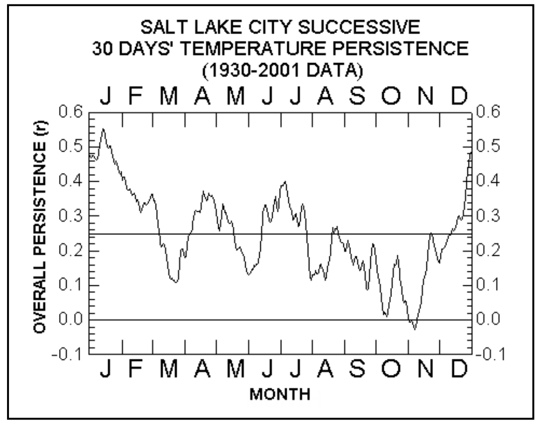

Figure 7. Successive 30 Days' Overall Temperature Persistence Statistics for Salt Lake City, Utah

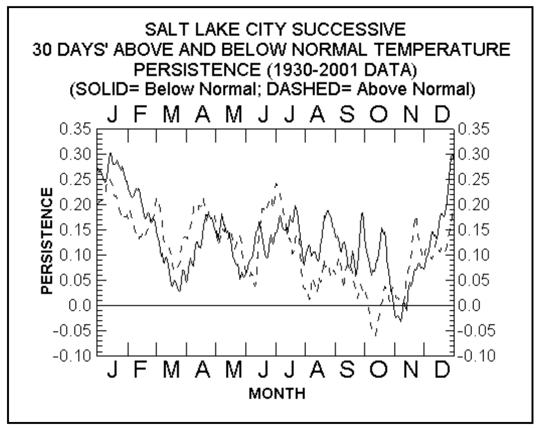

Figure 8. Successive 30 Days' Above and Below Normal Temperature Persistence Magnitudes for Salt Lake City, Utah

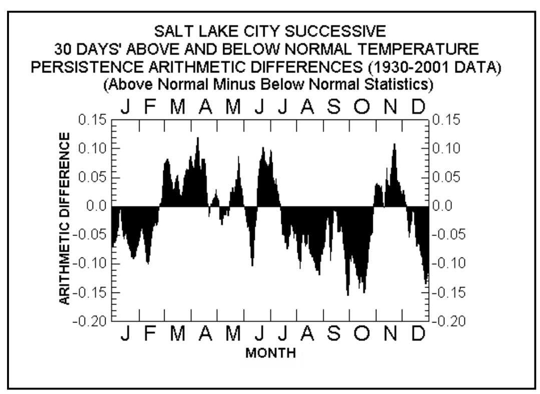

Figure 9. Successive 30 Days' Above and Below Normal Temperature Persistence Magnitude Differences for Salt Lake City, Utah

Examining Figure 9, cold persistence exceeds warm persistence for all sequence midpoints from early December through late-February, the converse being true from late February through early April. The most striking feature is the excess of cold over warm from the early July to early November midpoints. The below normal relative excess for the September 18 to October 17 vs. October 18 to November 16 sequence (-.149), the most pronounced, has an estimated t-value of –2.0, significant at the .02 level for a one-tail test.

7. SIXTY-DAY TEMPERATURE PERSISTENCE FOR LOS ANGELES CIVIC CENTER

Figures 10 to 12 trace the respective Overall Persistence, Above (Below) Normal Persistence, and Above/Below Normal Persistence difference statistics for the Los Angeles, California Civic Center, based on data from the start of 1921 through February 2001. According to the NCDC International Station Meteorological Climatic Summary, "The Pacific Ocean is the primary moderating influence", and at the station there is "a variable balance between the [westerly, onshore] sea-breezes and either hot or cold [offshore] winds from the interior resulting in some variety in weather conditions". The frequency mix between onshore and offshore flow together with sea-surface and Great Basin air mass temperature anomalies determine to a large extent the temperature persistence character during the year.

From Figure 10, Los Angeles Civic Center's average overall persistence (r=+.344) is the highest of the four stations. Two prominent relative maxima, centered in mid-June (r=+.546 for May 21 to June 19 vs. June 20 to July 19) and mid-August (r=+.564 for July 16 to August 14 vs. August 15 to September 13), respectively, are present. There is also a series of secondary peaks in the early part of the year through mid-May, and one in mid-November, the latter (r=+.411 for October 17 to November 15 vs. November 16 vs. December 15) more conspicuous, as fall and early winter magnitudes are generally lower than those for the other times of the year. Based on absolute, not relative magnitude, the only minimum of note (r=+.099 for the September 2 to October 1 vs. October 2 to October 31 sequence) likely reflects a shift from the predominant onshore regime of

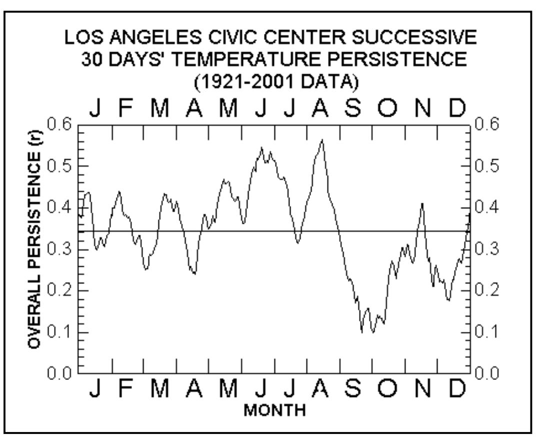

Figure 10. Successive 30 Days' Overall Temperature Persistence Magnitudes for Los Angeles Civic Center, California

summer to a more alternating onshore and offshore (e.g., Santa Ana episodes) pattern during October. Offshore flow becomes even more frequent in November and December.

Figure 11 shows the breakdown into above normal and below normal components. For the first half of the year, cold persistence shows a more even pattern than warm, the undulating pattern of the latter perhaps the influences of a relatively few individual years that displayed particularly prolonged and extreme above normal temperature regimes corresponding to those sequences. Highest individual cold persistence statistic (+.290) is noted for the July 14 to August 12 vs. August 13 to September 11 sequence, a secondary maximum (+.256) occurring for December 11 to January 9 vs. January 10 to February 8. Highest warm persistence figure (+.305) occurs for the June 20 to July 19 vs. July 20 vs. August 18 sequence, a secondary maximum (+.276) appears for the August 15 midpoint, just two days after the extreme maximum cold persistence figure. Both types show a rapid decline after mid-August, cold persistence actually exhibiting slightly negative figures (-.005) at the October 2 and 3 midpoints. Figure 11 indicates that the late November overall persistence peak is mostly cold persistence in makeup. Since offshore flow is rare on a synoptic-scale basis in mid-summer, the character and sign of the anomalies at this season most likely relate to sea-surface temperatures. However, in winter, it is likely a combination of both sea-surface temperature abnormalities and Great Basin air mass character. During some winters, intensely cold air masses get trapped in the Basin, their subsiding outflow spilling through the Southern California passes and canyons, creating potentially long spells of below normal temperatures near the coast.

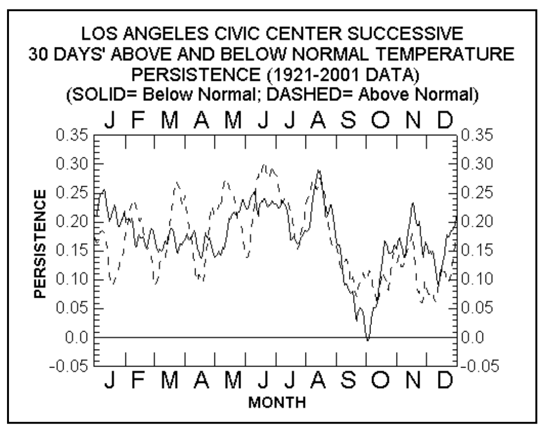

Figure 11. Successive 30 Days' Above Normal and Below Normal Temperature Persistence Magnitudes for Los Angeles Civic Center

From Figure 12, cold persistence predominates over warm persistence for almost every sequence from the early October through early February midpoints, warm persistence over cold for the great majority of points

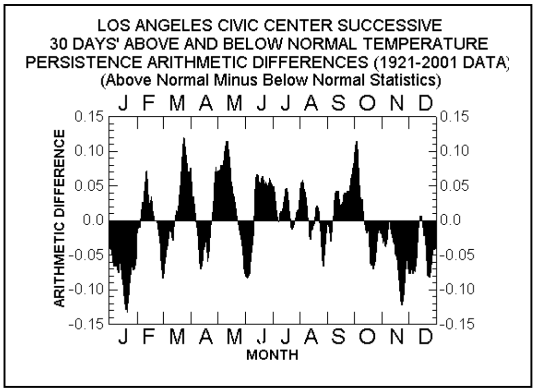

Figure 12. Successive 30 Days' Above and Below Normal Temperature Persistence Magnitude Differences for Los Angeles Civic Center

over the early June to early October interval. The largest discrepancy (-.132), reflecting the deficit of warm persistence (+.098) vs. cold (+.230), for the December 22 to January 20 vs. January 21 to February 19 sequence, has an estimated t-value of 1.9 and is significant at the .03 level. Greatest warm persistence excesses (around +.120 for sequences centered in late March, early May, and early October) are significant at the .05 level.

8. SUMMARY

Using standardized data and taking advantage of the statistical principle that defines the correlation coefficient as the sum of individual z-score products divided by the sample size (n), the annual course of successive 30-days' temperature persistence (overall, above normal, and below normal) was investigated at one-day intervals for four U.S. stations with lengthy periods of record: New York City Central Park, Minneapolis-St. Paul, Salt Lake City, and the Los Angeles Civic Center. Results showed a variety of overall persistence patterns, three of the stations showing summer absolute maxima, one with a winter maximum (Salt Lake City). Persistence minima, corresponding to natural transition periods or breakpoints, frequently occurred in the Spring and Fall.

By grouping appropriate leading series cross-products into subsets, summing them, and then dividing by n, overall persistence was mechanically "decomposed" into "above normal" (or "warm") and "below normal" (or "cold") components. In addition, arithmetic differences between the two components were tested for significance based on results from a simulation study of standardized data with built-in correlations and sample sizes. With the possible exception of early April to early June cold persistence being more intense than warm for New York Central Park, no conclusively significant differences were encountered between the two types.

9. CONCLUSION

Future work will investigate the actual effects of leading time-series subset frequency discrepancies (important or not) on standard error magnitudes like those in Table 1. Preliminary results, for example, on the n=150 and ρ=+.25 combination indicates that there is a slight increase in the standard errors as the subset discrepancy gets larger (overall standard error, from Table 1, is 0.050).

In the 5000 groups generated for this combination, median subset frequency discrepancy was 4 (e.g., 73 cases for one subgroup, 77 for the other). Standard error magnitudes for this discrepancy figure and lower (i.e., 4, 2, and 0) was 0.049. That for 6, 8, and 10 (corresponding to discrepancies above the median to about the upper decile) was 0.050; that for discrepancies at 12 and above was 0.052.

Thus, while there is an increase, it is small, and the introductory statement that subset frequency discrepancies, especially with regards to long periods of record, would likely have relatively unimportant effects on interpretation of warm/cold persistence magnitudes and the arithmetic differences between them is probably reasonable.

10. ACKNOWLEGEMENTS

Much of the data were downloaded from the Utah Climate Center online site: http://climate.usu.edu/free/default2.htm and the National Climatic Data Center site: http://www.ncdc.noaa.gov/

11. REFERENCES

Fisk, C.J., 1995: "A Technique for Determining the Statistical Significance of Warm Persistence/Cold Persistence Contrasts Within Decomposed Correlation Coefficient Temperature Persistence Measurements", *Proceedings of the 9th Conference on Applied Climatology, American Meteorological Society, Dallas.*

Johnson, R. A., and Wichern, D. W., 1982*; Applied Multivariate Statistical Analysis,* Prentice-Hall Inc., 594 pp.

NOAA.,1993; *International Station Meteorological Summary: Version 3.0*, National Climatic Data Center, Asheville,N.C. (CD-ROM).

Wilkinson, Leland, 1990: *SYSTAT: The System for Statistics,* SYSTAT, Inc., 677 pp.