P3.1 OBJECTIVE IDENTIFICATION OF EXTREMEMOST MIDNIGHT-TO-MIDNIGHT HOURLY HISTORICAL TEMPERATURE PATTERNS UTILIZING PRINCIPAL COMPONENTS ANALYSIS

Charles J. Fisk *
Naval Base Ventura County - Point Mugu, CA

1. INTRODUCTION

Hourly temperature readings are a standard meteorological element measured at first-order weather stations. Due perhaps to their sheer numbers (744 readings for a single 31-day calendar month, many thousands for a 30-year climatic "normals" period), their mundane character (usually), and the convenience of using the more concise summary-of-the day statistics like daily maxima and minima, detailed statistical analyses of hourly temperature pattern variations for the same station are less common. Diurnal temperature time-series, though, like their summary-of-the-day counterparts, should have interesting variations in pattern worthy of study. An array of twenty-four mean hourly climatological temperatures for a given calendar month, for example, can serve as that month's idealized "normal" diurnal pattern, most likely depicted by some form of sine wave or first harmonic, but extraction of the most anomalous patterns that have actually occurred over history by some objective statistical means should be interesting also, analogous to identifying record extreme daily maximum and minimum temperatures. A previous study along these conceptual lines [Fisk, 2004] investigated Downtown Los Angeles extreme patterns in summary-of-the-day maximum and minimum temperatures (depicted as floating-bars) back to 1921, comparing inter-year "configurations" between the same calendar month, and between complete years.

The same motivation is applied here involving a lengthy history of midnight-to-midnight "hourlies". Utilizing Los Angeles International Airport 1939 and 1947-2010 temperature data, downloaded and processed from the NCDC Integrated Surface Hourly ("ISH") online site, the utility of Principal Components Analysis (Correlation and Covariance, each unrotated) is demonstrated. First component correlation loadings characterize "shape", first covariance loadings, "spread". The highest and lowest correlation/covariance loadings identify the most anomalous patterns in terms of these attributes, and for a set of calendar months (January, April, July, and October), the most extreme patterns are identified and described. A few noteworthy examples from other months are also discussed, along with presentation of a graphical example identifying the most anomalous July patterns in a 2-D ("shape"/"spread") sense, employing Kernel Ellipsoid contouring methodology.

* *Corresponding author address:* Charles J. Fisk, Naval Base Ventura County, Point Mugu, CA. 93042: e-mail: charles.fisk@navy.mil

2. METHODOLOGY

The "shape" and "spread" concepts utilized here are borrowed from Yarnal [1993], who compared and contrasted the use of correlation and covariance PCA matrices in the context of synoptic climatology map pattern-recognition and classification. Applied to map interpretation, the correlation coefficient was found to be strongly influenced by contour feature or "shape" similarities, absolute magnitude contrasts or offsets across otherwise identical patterns producing identical correlations. Covariance, the formula of which is identical to that of the correlation coefficient except that the x and y standard deviation product term is excluded from the denominator, incorporated the absolute magnitude differences between patterns that the correlation is transparent to - a kind of "spread" measure.

Adapting these general "shape" and "spread" notions to a time-series application, Minneapolis-St. Paul, MN floating bar daily temperature patterns were analyzed by comparing average correlation and average covariance statistics of individual-year maximum/minimum temperature time-series vs. other years of the same chronology [Fisk, 1995]. For example, a given year's "average correlation" would be the mean correlation of its time-ordered temperature observations vs. all the other individual years' observations covering the same calendar segment. The most "anomalous" floating bar patterns were indicated by the highest or lowest mean correlation or covariance rankings. Highest average correlations were associated with day-to-day maximum/minimum patterns that closely resembled climatological maxima and minima profiles ("shape"), irregardless of overall departure from normal. Average covariances ("spread") were affected by atypical temperature progressions, extreme values, peculiar patterns in average daily ranges, and unusually pronounced seasonal contrasts.

In a subsequent study on floating-bar temperature patterns [Fisk, 2004] which directly utilized the linear PCA matrix concepts described by Yarnal, Downtown Los Angeles daily max/min temperature patterns were examined back to 1921. Along the way, it was discovered that the unrotated first component correlation and covariance loadings' statistics, respectively, were exact analogs of the average correlation and average covariance statistics generated in the earlier study - different in magnitude but perfectly correlated, linearly. Moreover, correlation PCA first component standardized scores were perfectly correlated with their counterpart climatological mean daily maxima and minima statistics. Thus, a given year's correlation loading figure (or "shape" measure) was equivalent to the correlation coefficient of the temperature array in question versus reference climatology. In like fashion, the covariance loading (or "spread" metric) was equivalent to the covariance of the temperature array vs. climatology. In calculating the correlation and covariance loadings statistics, the desktop software used standardized climatology scores, the resulting covariance loadings statistics much smaller in magnitude than otherwise and thus of a closer order of magnitude to those of the correlation loadings. The correlation loadings were unaffected by scores' standardization.

Terminological equivalencies aside, the powerful PCA matrix-reducing software was a speedy and advantageous means of generating and ranking the required "shape" and "spread" statistics, be they labeled "first component unrotated correlation/covariance loadings", or the terminologically more familiar but equivalent "linear correlation/covariance coefficients" .

In this particular application, desktop software matrix manipulating power was especially useful, as sample sizes approached 2000 cases in some instances.

3. DATA AND PROCEDURES

Generation of pattern statistics for a calendar month of interest was accomplished by first arranging the historical data in rectangular arrays of n columns ("n" number of cases or individual days with complete hourly observation sets) and r rows ("r" being the number of hours in a day: 24). For example, the month of July had 1946 complete diurnal profiles to examine, the resulting setup matrix comprising 24 rows and 1946 columns.

Correlation and Covariance PCA's were then performed on the arrays (O-mode decomposition with no rotation), the loadings' statistics compiled by day and ranked by magnitude.

4. RESULTS

4.1 - *January Diurnal Results*

Figure 1 (From Table 1) is a graph of those diurnal temperature series that produced the most extreme January Correlation and Covariance loadings statistics (four curves plus climatology). Some 1885 cases were analyzed, about 93.5% of possible.

Inspecting Figure 1, the black "Mean Hourly" or climatological curve indicates that during January, LAX mean temperatures vary from around 50 F over the hours just before sunrise, to about 63 F just after Noon LST. The relative earliness of the peak reflects the sea-breeze, the LAX station just several miles east of the Pacific Ocean.

Examining the Correlation loadings' extremes in Table 1, a +.995 maximum statistic was generated for 21 January 1978 (blue trace in Figure 1), this particular day's temperatures essentially matching climatology through the early daylight hours, but with a slightly more amplified (warmer), but still symmetric pattern thereafter. Many other January profiles generated loadings figures above +.950, this being a typical pattern for LAX daily temperatures in January, not "extreme" in any real physical sense.

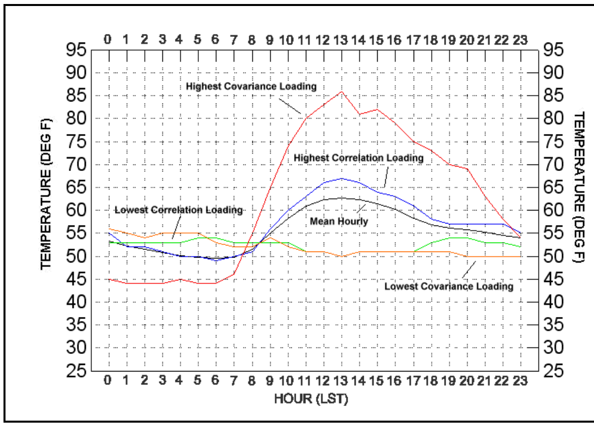

Figure 1 - Graph of LAX January diurnal temperature patterns producing the most extreme Correlation and Covariance Loading Statistics, along with climatological trace (From Table 1).

HRLST	CLIMATOLOGY	CLIMATOLOGY (STANDARD-IZED)	HIGHEST CORR LOADING (+.995) (21 JAN '78)	LOWEST CORR LOADING (-.795) (7 JAN '74)	HIGHEST COV LOADING (+15.02) (31 JAN '54)	LOWEST COV LOADING (-1.236) (21 JAN '39)
0	53.3 F	-0.51z	55 F	53 F	45 F	56 F
1	52.3	-0.73	52	53	44	55
2	51.5	-0.91	52	53	44	54
3	50.8	-1.06	51	53	44	55
4	50.2	-1.21	50	53	45	55
5	49.8	-1.30	50	54	44	55
6	49.5	-1.36	49	54	44	53
7	49.8	-1.29	50	53	46	52
8	51.6	-0.89	51	53	55	52
9	55.0	-0.11	56	53	65	54
10	58.3	0.62	60	53	74	52
11	61.0	1.22	63	51	80	51
12	62.4	1.55	66	51	83	51
13	62.7	1.62	67	50	86	50
14	62.3	1.53	66	51	81	51
15	61.5	1.34	64	51	82	51
16	60.2	1.05	63	51	79	51
17	58.2	0.59	61	51	75	51
18	56.8	0.28	58	53	73	51
19	56.2	0.14	57	54	70	51
20	55.8	0.05	57	54	69	50
21	55.3	-0.06	57	53	63	50
22	54.6	-0.21	57	53	58	50
23	54.0	-0.34	55	52	54	50

Table 1 – Tabular Version of January Hourly Temperature Time-series plotted in Figure 1, along with standardized climatological series (utilized in covariance loadings' calculations)

As the January hourly climatological profile is a first harmonic function (i.e., the diurnal cycle), the high proportion of individual series with loadings exceeding +950 reflects the fact that the vast majority of hourly temperature patterns at LAX in January conform to that cycle, and among other things, the distribution of correlation loadings is likely very skewed.

To reiterate in non-PCA terms, the +.995 correlation loading (or "shape") figure of the 21 January 1978 diurnal series (column 4) was simply the correlation coefficient of its' array with climatology (either column 2 or column 3) .

Lowest Correlation loading figure (-.795) was produced for 7 January 1974 (green trace in Figure 1, and column 5 in Table 1), a day with a slightly “anti-diurnal” shape (coldest temperatures of the day, – either 50 F or 51 F, noted over the late morning and afternoon hours, 1100 to 1700 LST). The 50 F low reading was recorded at 1300 LST, the very hour of the day with the highest climatological mean (62.7 F – see column 2). Not surprisingly, 7 January 1974 was cloudy and wet, rain noted for 23 of the 24 observations, but the slightly “inverted” positioning of the day’s coldest temperatures in the late morning and afternoon set this pattern apart in the “shape” sense.

Highest covariance figure (+15.02) was generated for 31 January 1954 (red trace in Figure 1, and column 6 in Table 1) an “offshore flow” or “Santa Ana” day, with hour-to-hour temperatures soaring more than 40 F, from 44 F just before sunrise to 86 F at 1300 LST. The high temperature range (and low overnight readings, significantly colder than climatology) suggests that the Santa Ana episode did not set in until after sunrise. The delayed onset and pronounced temperature rise when it did begin distinguishes this particular day’s pattern in terms of “spread”.

In non-PCA terms, the +15.02 covariance loading statistic was the covariance of the 31 January 1954 series (column 6 in Table 1) with the *standardized* hourly climatological means in column 2 of the Table. This figure was also the highest such statistic for any month (January-December) in the 65-year history.

Lowest covariance loading (-1.236) was for 21 January 1939 (orange trace in Figure 1, column 7 in Table 1). On this day, temperatures exhibited a slow declining trend, from 56 F at midnight to 50 F by 2000 LST, again attributable to the evaporative cooling effects of a rainy, cloudy day, to go with a cold frontal passage. Rain was recorded for 22 of the 24 hourly observations.

It’s obvious that the 21 January 1939 curve is similar to that of 7 January 1974 (both displaying very atypical “shape” and “spread”). Indeed, the latter’s covariance loading statistic (-0.935) was the third lowest of the history.

4.2 - *April Diurnal Results*

Figure 2 (From Table 2) is a graph of those actual diurnal temperature observations that produced the most extreme Correlation and Covariance loadings statistics (four curves plus climatology) for April. There were 1896 cases available, about 97.2 % of possible.

The black “Mean Hourly” curve shows a diurnal range from about 54 F at 0500 LST to 65 F for 1100 LST to 1400 LST, inclusive (see also column 2 in Table 2 – column 3 is the standardized version of column 2).

Highest correlation loading figure (+.994) was generated for 29 April 1977 (blue trace and column 4), a few degrees warmer than normal overall, but with a very smooth diurnal rise and fall in hour-to-hour readings In this regard, it may be recalled that the correlation loading (or coefficient) is a purely “shape” metric, transparent to the pattern’s overall departure from normal. However, the magnitude of the statistic here was enhanced slightly by the 29 April 1977’s amplified readings versus climatology for the afternoon hours, the correlation coefficient being sensitive to extreme values (in this case the day’s warmest ones matching up well with climatology’s highest). This same effect was seen for the 21 January 1978 case.

Lowest Correlation Loading (-.293) was produced for 29 April 1987 (green trace and column 5), a day with another slightly “anti-diurnal” pattern, afternoon temperatures mostly colder than the morning ones.

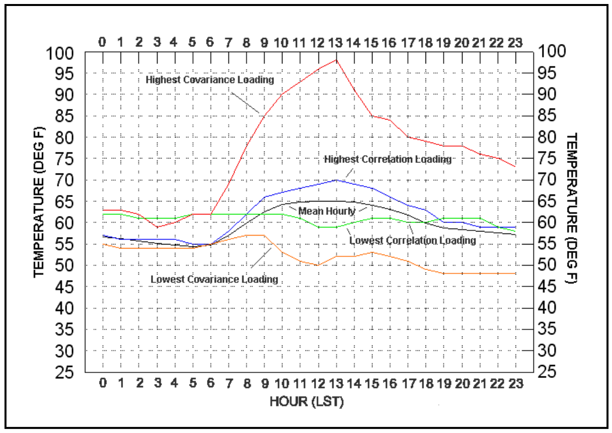

Figure 2 - Graph of LAX April diurnal temperature patterns producing the most extreme Correlation and Covariance Loading Statistics, along with climatological trace (From Table 2).

HRLST	CLIMATOLOGY	STANDARD-IZED CLIMATOLOGY	HIGHEST CORR LOADING (+.994) (29 APR '77)	LOWEST CORR LOADING (-.293) (29 APR '87)	HIGHEST COV LOADING (+11.69) (4 APR '89)	LOWEST COV LOADING (-.452) (20 APR '57)
0	56.7 F	-0.76 z	57 F	62 F	63 F	55 F
1	56.2	-0.88	56	62	63	54
2	55.6	-1.03	56	61	62	54
3	55.1	-1.17	56	61	59	54
4	54.7	-1.27	56	61	60	54
5	54.4	-1.36	55	62	62	54
6	54.7	-1.27	55	62	62	55
7	57.0	-0.68	58	62	69	56
8	59.8	0.07	62	62	78	57
9	62.5	0.77	66	62	85	57
10	64.2	1.22	67	62	90	53
11	64.9	1.40	68	61	93	51
12	65.1	1.45	69	59	96	50
13	65.1	1.45	70	59	98	52
14	64.7	1.36	69	60	91	52
15	64.1	1.18	68	61	85	53
16	63.1	0.93	66	61	84	52
17	61.7	0.57	64	60	80	51
18	59.9	0.10	63	60	79	49
19	58.7	-0.23	60	61	78	48
20	58.3	-0.34	60	61	78	48
21	58.0	-0.42	59	61	76	48
22	57.7	-0.49	59	59	75	48
23	57.2	-0.61	59	58	73	48

Table 2 – Tabular Version of April Hourly Temperature Time-series plotted in Figure 2, along with standardized climatological series (utilized in covariance loadings’ calculations)

Readings were either 61 F or 62 F over midnight to Noon LST, but between 58 F and 61 F thereafter.

Inspecting the "spread" statistics, highest covariance loading (+11.69) was for 4 April 1989 (red trace and column 6), a unseasonably warm offshore flow day in early Spring in which the mercury climbed from 59 F at 0300 LST to 98 F by 1300 LST, some 33 F warmer than the April norm for the hour.

Lowest covariance loading (-.452) was created for 20 April 1957 (orange trace and column 7), a day displaying an irregular but downward trending pattern. The day's warmest readings came at 0800 and 0900 LST (57 F), the mercury falling to 50 F by Noon LST, recovering to 53 F by 1500 LST, then declining again to 48 F by 1900 LST, remaining at that level through 2300 LST.

4.3 - July Diurnal Results

Figure 3 (From Table 3) is a graph of the LAX July hourly profiles that resulted in the most extreme Correlation and Covariance loadings statistics, along with climatology. Some 1946 cases were analyzed, 96.6% of possible.

The "Mean Hourly" curve (black trace in Figure 3 derived from column 2 in Table 3) shows a diurnal range from around 64 F for the hours 0200 to 0500 LST to about 73 F for 1100 to 1400 LST. This daily spread, even smaller than that for January and April, again reflects the cooling sea-breeze effect. Stations located farther inland would typically have their summertime daily maxima more pronounced and a few hours later in the afternoon.

Highest July correlation loading ("shape") statistic was +.995, recorded for 19 July 1977 (blue trace and column 4). Like its January and April counterpart curves for this particular extreme, 19 July 1977 exhibits a very smooth hour-to-hour progression. In contrast with the January and April profiles, however, its temperature pattern is amplified slightly in the below normal direction for the early morning and evening hours, conforming closely to climatology for the afternoon ones

Lowest correlation loading figure (+.506) was produced for 28 July 1972 (green trace and column 5), a day with a genuinely anomalous pattern in the physical sense. Maximum temperature for the day (84 F) came at 2300 LST, courtesy of an unseasonable offshore/ downslope flow event. Temperatures further inland had been very warm during the day, Downtown Los Angeles recording over 100 F, LAX reaching 79 F at 1100 and 1200 LST, but a peculiar set of meteorological circumstances brought even warmer temperatures to LAX well after sunset, following a cooling to 74 F by 1900 and 2000 LST. Winds were NNW to NNE at 5 to 8 knots during the peak three hours of the temperature rise.

Maximum Covariance Loading ("spread") statistic (+8.413) was generated for 4 July 1957 (red trace and column 6). Temperatures soared from 65 F just after sunrise to 89 F by 1000 LST, retreated in response to the sea-breeze to 85 F for 1100 and Noon LST, then surged back to 92 F at 1300 LST.

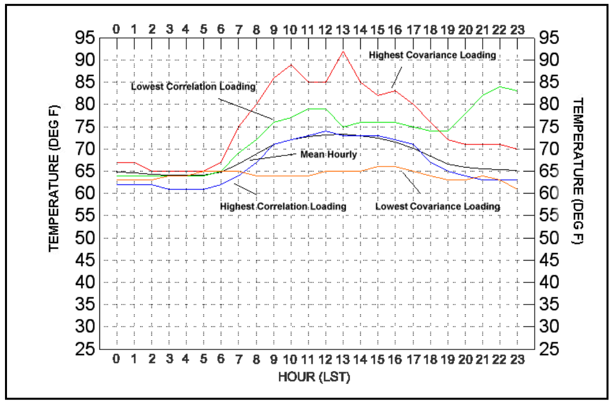

Figure 3 - Graph of LAX July diurnal temperature patterns producing the most extreme Correlation and Covariance Loading Statistics, along with climatological trace (From Table 3).

HRLST	CLIMATOLOGY	STANDARD-IZED CLIMATOLOGY	HIGHEST CORR LOADING (+.995) (19 JUL '77)	LOWEST CORR LOADING (+.506) (28 JUL '72)	HIGHEST COV LOADING (+8.413) (4 JUL '57)	LOWEST COV LOADING (+.615) (25 JUL '65)
0	64.8	-0.92	62	64	67	63
1	64.6	-0.97	62	64	67	63
2	64.3	-1.04	62	64	65	63
3	64.2	-1.08	61	64	65	64
4	64.1	-1.11	61	64	65	64
5	64.1	-1.12	61	64	65	65
6	64.8	-0.92	62	65	67	65
7	66.5	-0.42	64	69	75	65
8	68.8	0.23	67	72	80	64
9	71.0	0.83	71	76	86	64
10	72.1	1.16	72	77	89	64
11	72.8	1.35	73	79	85	64
12	73.2	1.46	74	79	85	65
13	73.3	1.49	73	75	92	65
14	73.0	1.42	73	76	85	65
15	72.4	1.25	73	76	82	66
16	71.5	0.98	72	76	83	66
17	70.1	0.58	71	75	80	65
18	68.4	0.10	67	74	76	64
19	66.6	-0.39	65	74	72	63
20	65.8	-0.61	64	78	71	63
21	65.5	-0.70	63	82	71	64
22	65.3	-0.76	63	84	71	63
23	65.1	-0.82	63	83	70	61

Table 3 – Tabular Version of July Hourly Temperature Time-series plotted in Figure 3, along with standardized climatological series (utilized in covariance loadings' calculations)

Minimum Covariance Loading figure (+0.615) was produced for 25 July 1965 (orange trace and column 7), a day with readings hovering around 65 F plus or minus a few degrees throughout.

4.4 - *October Diurnal Results*

Figure 4 (From Table 4) is a graph of the LAX October hourly profiles that produced the most extreme Correlation and Covariance loadings statistics, along with climatology. Some 1930 cases were analyzed, 95.8% of possible.

The “Mean Hourly” curve (black trace in Figure 4 derived from column 2 in Table 4) shows a diurnal range from around 60 F for the hours 0500 and 0600 LST to about 72 F at Noon LST.

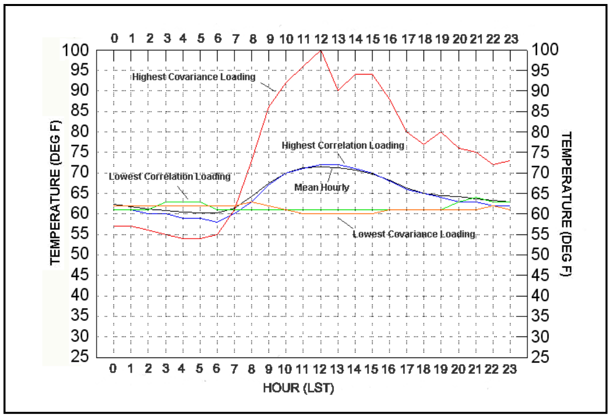

Figure 4 - Graph of LAX October diurnal temperature patterns producing the most extreme Correlation and Covariance Loading Statistics, along with climatological trace (From Table 4).

HRLST	CLIMATOLOGY	CLIMATOLOGY (STANDARD-IZED)	HIGHEST CORR LOADING (+.997) (4 OCT '57)	LOWEST CORR LOADING (-.456) (24 OCT '05)	HIGHEST COV LOADING (+14.97) (3 OCT '53)	LOWEST COV LOADING (-.688) (31 OCT '95)
0	62.4 F	-0.69 z	61 F	61 F	57 F	62 F
1	61.8	-0.83	61	61	57	62
2	61.3	-0.98	60	61	56	62
3	60.8	-1.10	60	63	55	62
4	60.5	-1.18	59	63	54	62
5	60.2	-1.26	59	63	54	62
6	60.2	-1.25	58	61	55	62
7	61.4	-0.93	60	61	61	62
8	64.3	-0.20	63	61	73	63
9	67.4	0.58	67	61	86	62
10	70.0	1.24	70	61	92	61
11	71.4	1.58	71	61	96	60
12	71.6	1.63	72	61	100	60
13	71.3	1.57	72	61	90	60
14	70.6	1.40	71	61	94	60
15	69.7	1.15	70	61	94	60
16	68.2	0.79	68	61	88	61
17	66.2	0.29	66	61	80	61
18	64.9	-0.05	65	61	77	61
19	64.5	-0.15	64	61	80	61
20	64.2	-0.24	63	63	76	61
21	63.7	-0.35	63	64	75	61
22	63.4	-0.44	62	63	72	62
23	62.9	-0.57	62	63	73	61

Table 4 – Tabular Version of October Hourly Temperature Time-series plotted in Figure 4, along with standardized climatological series (utilized in covariance loadings’ calculations)

Highest October correlation loading figure was +.997 (also the highest for any month), generated by the diurnal pattern of 4 October 1957 (blue trace in Figure 4, column 3 in Table 4). The hourly profile is very smooth, and based on the previous results of January, April, and July, presumably a necessary condition for a high-end statistic of this kind. It also conforms very closely to climatology, with only a slight indication of a July-like below-normal amplification for the evening and early morning hours.

Lowest Correlation Loading (-.456) is for 24 October 2005, (green trace and column 5), a day that had steady 61 F temperatures for fourteen straight hours from 0600 to 1900 LST, inclusive, followed by a slight warming to 64 F at 2100 LST, the day’s highest. .

Highest covariance loading statistic (+14.97, second only to 31 January 1954 in high magnitude – red trace and column 6) was produced for 3 October 1953, a major offshore flow event. This day saw morning readings before sunrise in the mid to high 50’s, 5 F or more colder than climatology, but from 0700 LST on, temperatures climbed rapidly, to 100 F at Noon LST. A sea-breeze dropped the mercury to 90 F next hour, but 94 F was observed again at 1400 and 1500 LST.

Lowest covariance statistic (-.688) was generated for 31 October 1995 (orange trace and column 7), a day with the lowest temperatures (60 F) coming over a five hour interval from 1100 to 1500 LST. All the others of the day except for a 63 F at 0800 LST were either 61 F or 62 F.

4.5 – Other Extreme Patterns of Interest.

While the four sample months examined above have provided a sense of what extreme pattern features are captured by which extreme statistics, Table 5 lists a few other anomalous hourly profiles of interest that encompassed the other eight months.

The highest shape magnitude pattern has already been presented (4 October 1957) as well as the highest and second highest spread ones (31 January 1954 and 3 October 1953, respectively), and Table 5 lists those associated with the lowest and second shape and spread metrics, respectively, for any month.

From the Table, interestingly, 16 Dec 1987 has both the lowest (or most negative) spread and shape statistics (-2.501 and -.832, respectively – columns 2 and 3) for any day in the 65-year history, its overall pattern a combination of the downward trending and “anti-diurnal” features that characterize low values of each of these metrics individually. Lowest readings of the day (43 F and 44 F) ran completely contrary to the normal diurnal cycle, coming over the peak daylight hours - 1000 to 1600 LST, respectively, and while a warming trend ensued for the evening hours before midnight (the mercury reaching the high 40’s), readings were still below pre-sunrise levels (52 F over 0200 to 0500 LST, inclusive); thus, producing a downward overall statistical trend in temperature also.

Second lowest spread statistic (-2.407) was generated for 5 Feb 1978 (column 4). A relatively steep overall temperature decline is evident from midnight to midnight (66 F to 58 F) and there is a modest anti-diurnal effect (56 F at 1000 and 1100 LST, 57 F from Noon to 1700 LST, 58 F thereafter) but the overall spread magnitude is still a bit higher than 16 Dec 1987’s.

HRLST	LOWEST COV LOADING FOR ANY MONTH (-2.501) (16 DEC '87)	LOWEST CORR LOADING FOR ANY MONTH (-.832) (16 DEC '87)	2ND LOWEST COV LOADING FOR ANY MONTH (-2.407) (5 FEB '78)	2ND LOWEST CORR LOADING FOR ANY MONTH (-.796) (3 FEB '98)
0	51 F	51 F	65 F	59 F
1	50	50	67	58
2	52	52	66	58
3	52	52	65	59
4	52	52	64	59
5	52	52	63	60
6	47	47	60	58
7	47	47	59	57
8	47	47	58	55
9	46	46	57	53
10	44	44	56	54
11	44	44	56	53
12	43	43	57	53
13	44	44	57	53
14	44	44	57	52
15	44	44	57	52
16	44	44	57	54
17	46	46	57	54
18	46	46	58	52
19	46	46	58	52
20	46	46	58	53
21	48	48	58	53
22	49	49	58	54
23	49	49	58	54

Table 5 - A few other Extreme Patterns of Note for those months other than January, April, July, and October.

.

Second lowest shape statistic (-.796), was produced for 3 February 1998. It shows another "anti-diurnal" or slightly "bowl"-like pattern, with the lowest readings of the day (52 F and 53 F) scattered mostly for the afternoon hours.

5. JOINT CORRELATION ("SHAPE") AND COVARIANCE ("SPREAD") LOADINGS' INFORMATION

With two metrics available to characterize hourly temperature patterns, one could take this to another level by combining the two into a two-dimensional scatterplot analysis. Such was done in the 2004 study on the floating-bar patterns, utilizing bivariate normal confidence ellipsoids to identify and characterize the most extreme outliers. The calendar year shape and spread statistics were both univariate and bivariate normal, permitting straightforward analysis and plotting, but the sample individual months required a folded-log transformation on the shape (correlation) statistics to render them useable for such treatment.

The same bivariate normality 2-D approach was attempted here on the month of July's data, but the raw statistics, particularly the correlation loadings (again), were highly non-normal with no meaningful transformation possible. This of course was attributable to the fact that most July diurnal temperature profiles adhere closely to the diurnal cycle, resulting in the vast majority of loadings figures well above +.90, with a resulting highly asymmetric distribution.

To illustrate, Figure 5 and 6 are plots of the distribution of July correlation and covariance loadings distributions respectively. Clearly obvious is the skewed nature of the July shape metrics (mean correlation +.936). The spread distribution is more bell-shaped, but a scatterplot incorporating both these distributions would not be bivariate normal. Such a confidence ellipsoid approach is thus not feasible, but it is possible to construct contours utilizing distribution-free kernel ellipsoids.

Figure 7 is a scatterplot with the combined shape and spread statistics (n=1946) fitted to a p=.999 confidence kernel ellipsoid. In addition to a single large asymmetrically bounded cloud of points, there are also three smaller "islands". This particular p-level accomplished the objective of isolating a single outlier point, which could be considered the most anomalous in the 2-D sense. The points associated with the univariate extremes listed in Table 3 are identified with their dates labeled, and the most extreme 2-D point is identified in red. The 2-D case in question is the profile for 29 July 1972, the very day after the 28 July late evening temperature surge which produced the extreme lowest July ("anti-diurnal") spread statistic. The two successive days' profiles appear in Table 6. While the 29th's shape statistic (+.592) is higher than the 28th's (+.506), its spread magnitude (2.841) is a little lower than the 28th's (3.382), neither of these latter two especially anomalous, though. In this regard, however, the kernel algorithm evidently "considered" the 29th's relatively isolated position in the chart to be more extreme than the 28th's. The 28th's point, however, actually covers the p=.999 contour partially, a distinction not visible for any of the others, so from this property, it can be "designated" as the second most anomalous case. Overall, the shape and spread statistics in the chart are linearly correlated (r=+.441).

The late evening temperature rise event of the 28th played itself out over the early hours of the 29th (see Table 6), 83 F and 82 F observed at 0000 and 0100 LST, respectively, (18 F and 17 F above normal for those hours), but a sharp 11 F cooling set in at 0200 LST, readings from 0300 to 0600 LST all 69 F. The temperature started a strong advance over the morning hours, 85 F reached at 1100 AM (about 12 F above normal for the hour) but a sea-breeze brought another steep drop of 10 F to 75 F at Noon LST, hourly readings not exceeding 80 F again over the rest of the day.

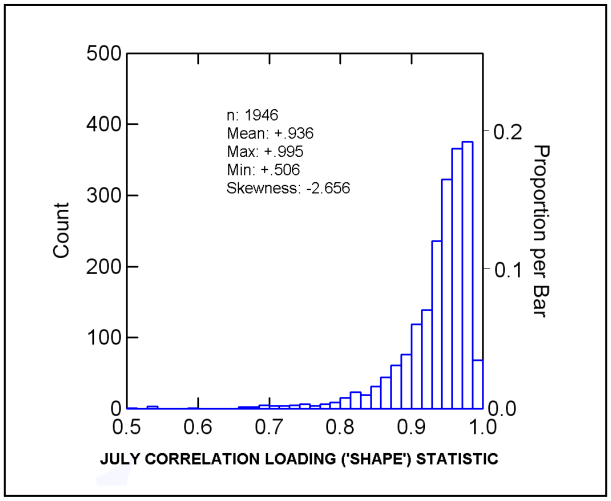

Figure 5 - Histogram of LAX July Correlation Loading ("Shape") Statistics

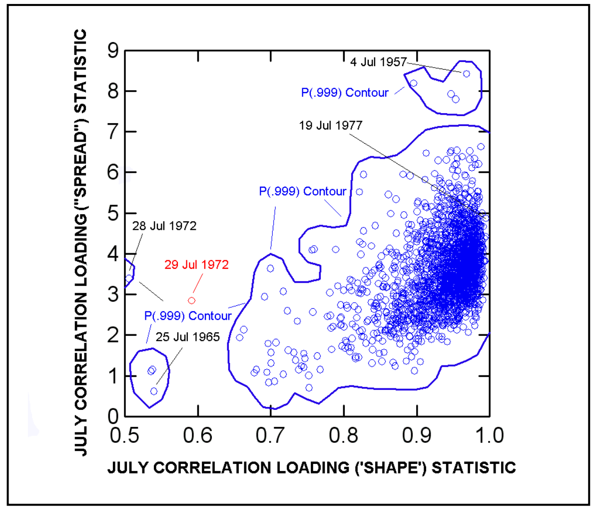

Figure 7 – 99.9% Confidence Kernel Contours for LAX July Joint Shape and Spread Statistics

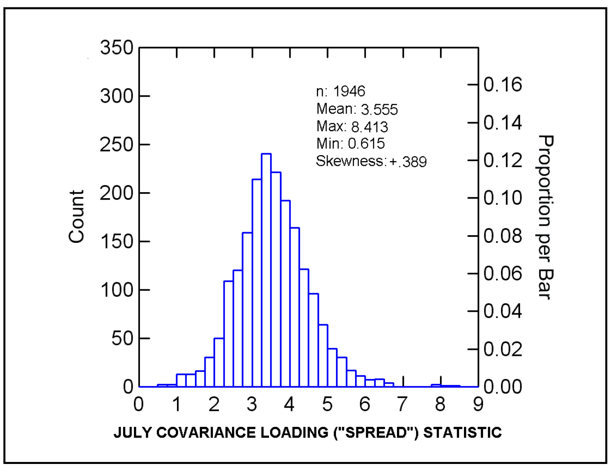

Figure 6 - Histogram of LAX July Covariance Loading ("Spread") Statistics

HRLST	CLIMATOLOGY	STANDARD-IZED CLIMATOLOGY	LOWEST CORR LOADING (+.506) (28 JUL '72)	MOST ANOMALOUS COMBINED 2-D PATTERN (29 JUL '72)
0	65 F	-0.92 z	64 F	83 F
1	65	-0.97	64	82
2	64	-1.04	64	71
3	64	-1.08	64	69
4	64	-1.11	64	69
5	64	-1.12	64	69
6	65	-0.92	65	69
7	67	-0.42	69	72
8	69	0.23	72	72
9	71	0.83	76	74
10	72	1.16	77	78
11	73	1.35	79	85
12	73	1.46	79	75
13	73	1.49	75	80
14	73	1.42	76	78
15	72	1.25	76	79
16	72	0.98	76	79
17	70	0.58	75	79
18	68	0.10	74	75
19	67	-0.39	74	73
20	66	-0.61	78	72
21	66	-0.70	82	72
22	65	-0.76	84	72
23	65	-0.82	83	71

Table 6: Hourly LAX Temperatures for 28-29 July 1972

6. SUMMARY AND CONCLUSIONS

Utilizing the concepts of shape and spread as originally applied to synoptic climatology map interpretation, the foregoing utilized unrotated first component correlation and covariance loadings statistics to identify the most anomalous diurnal temperature patterns that have been experienced at Los Angeles International Airport (LAX) for a selection of months over the 1939 and 1947-2010 period of record.

Highest correlation loading ("shape") statistics were represented by profiles that showed exceptionally smooth hour-to-hour progressions conforming closely to climatology (i.e., the diurnal cycle). Relatively low (sometimes negative) statistics described profiles that departed significantly from this – sometimes "anti-diurnal" in character, in which the coolest readings of the day came in the afternoon.

In contrast, highest covariance loading ("spread") statistics were seen for those profiles that experienced unusually amplified diurnal temperature rises and falls, in particular, days with cooler than normal readings for the pre-daylight hours, but pronounced temperature rises thereafter (as in a Santa Ana event). Finally, lowest spread statistics seemed to be confined to those profiles that exhibited a significant downward trend in temperature from midnight to midnight. As the extremely low-end shape and spread statistics were similarly anti-diurnal, they were often similar in rank, indeed, sometimes one and the same, as in the case of 16 Dec 1987 (see Table 5).

It should be said that since only the most extreme high or low-end statistics were evaluated here for a single station, other diurnal pattern scenarios for other stations are likely possible that would create similar extreme statistics. For example, patterns that displayed a consistent rise in temperature from midnight to midnight would likely generate extreme "shape" and "spread" statistics, although these types weren't encountered in the samples from the LAX data base.

Finally, it should be emphasized that analyzing diurnal profiles in terms of "shape" and "spread" is certainly not an exhaustive treatment of the subject. As a hypothetical example, a day that had the warmest hour-to-hour average in history, yet displayed an unremarkable diurnal shape and daily spread, would likely not produce extraordinary statistics of these two kinds – yet its pattern, by virtue of its record mean warmth, was certainly an anomaly in its own right. One alternative technique that might incorporate shape, spread and mean daily temperature effectively would be multivariate cluster analysis.

7. REFERENCES

Fisk, C.J., 2004: "Objective Identification of Extreme-Most Anomalous Daily Max/Min temperature Patterns using Principal Components Analysis", 17th Conference on Probability and Statistics in the Atmospheric Sciences, American Meteorological Society – Seattle, 2004. http:/ams.comfex.com/ams/pdfpapers/69198.pdf

Fisk, C.J., "An Objective Technique for Identifying Extreme-Most Pattern Anomalies in Daily Maximum/Minimum Temperature Time-Series", Proceedings of the 9th Conference on Applied Climatology, American Meteorological Society – Dallas, 1995

Yarnal, B., 1993: Synoptic *Climatology in Environmental Analysis, A Primer*, Bellhaven Press, 195 pp.

.