IDENTIFICATION AND ANALYSIS OF CONTIGUOUS HOURS' CLIMATOLOGICAL WIND PATTERN MODES UTILIZING K-MEANS CLUSTERING ANALYSIS COMBINED WITH THE V-FOLD CROSS VALIDATION ALGORITHM

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

1. INTRODUCTION

Climatological wind variability is an important meteorological element to be considered in planning, forecasting, and decision-making activities in which wind conditions are crucial on some level. Wind rose diagrams, for example, can provide insights into the wind character for individual hours of interest by depicting the most favored compass directions and associated speeds. Resultant wind calculations can be valuable in producing distilled single-value statistics derived from many different individual observations.

Also operationally useful but more complicated to produce would be characterizations of a station's most important *contiguous* hour-to-hour wind patterns, encompassing, for example, an entire day, midnight-to-midnight. In the same manner as there are favored individual hourly directions and related speeds, there are undoubtedly preferred, adjacent hour-to-hour patterns, or "modes". Such patterns are also likely to exhibit varying seasonal inclinations. This idea could be extended further, considering multiple stations' simultaneous hourly wind patterns.

Resolving' idealized wind patterns of this kind can be accomplished by a clustering analysis, making use of individual hour wind observations' u and v components, and to this end, K-Means Clustering Analysis integrated with a special add-on capability, the V-Fold Cross-Validation Algorithm, is employed.

Traditional K-Means being a trial-and-error procedure, the V-Fold Algorithm is an automated, iterative training sample procedure that rapidly produces in ascending order, 2 to K cluster sets, the iterations ceasing at some "optimal" number K, depending on a choice of statistical distance metric (Euclidean, Squared Euclidean, etc.), percent improvement cutoff threshold, and other settings. The software that offers the K-Means/V-Fold capability, however, also gives the option of fixing the number of clusters, and this study, while still utilzing the V-fold algorithms' processing speed and efficiency, bypasses the cutoff feature. The number of wind pattern clusters to be generated is arbitrarily fixed at six, the rationale stemming from prior experience and experimentation, which determined that lesser modes, normally "cutoff", still conveyed useful climatological information. The fixed cluster approach has already been applied with good results on many stations' data in the Southern California area, both for operational and purely informative purposes, including those of several NOAA Buoys in the Southern California Bight area, local military bases/NWS Offices and automatic weather stations As a demonstration of this technique, a set of six midnight-to-midnight hourly wind pattern clusters are produced for the Point Mugu, CA station, a naval base near Oxnard, using the squared Euclidean distance metric. In addition to the mean vector wind representations, output includes percent frequency comparisons, month-by-month. Following this is a similar two-station analysis incorporating the California High Desert stations of Daggett and Victorville.

2. DATA AND PROCEDURES

Period of record for the Pt. Mugu station was 1963-2015, that for the joint Daggett/Victorville data sets 1951-2014. The raw hourly data were downloaded, decoded, QC'd, and processed from the Integrated Surface Hourly ("ISH") data base, accessible from the National Climatic Data Center online site ("NCDC-online"). Only those individual days that had complete hour-to-hour wind observation sets were retained.

The STATISTICA Data Miner Clustering module was utilized to employ this technique. Preliminary to the analyses, the Pt. Mugu, Daggett, and Victorville data were were normalized (an automatic software feature) to reduce them to a common scale and lessen the influence of outliers.

For a single station midnight-to-midnight analysis, the clustering analysis would be preceded by decomposition of the hourly observations into their respective u and v components; the analysis thus being one in 48-dimensional space (the two-station analysis would be a 96-dimensional application). Each of the resulting six individual cluster "clouds" would have midpoints or centroids in 48-D space, made up of mean u and v values for each hour of the day. These would then be recombined, hour-by-hour using the arctangent function, producing 24 sets of hourly mean vector wind statistics; The cluster arrays would be independent self-contained entities, depicting idealized progressions of mean vector wind character hour-by-hour over the course of a day.

Since the clustering process assigns diurnal wind observations of a very similar character to a given group, the individual hour mean vector wind direction results in many cases could be interpreted, with only slight loss of generality, as *average wind directions.* Also, the accompanying mean vector wind *speeds*, would be very similar (slightly less) than their counterpart mean scalar wind speed statistics,

A "one-size-fits-all' approach was taken with the clustering analysis applied to all months' data as a

----

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

single unit. A validation of this tactic, also based on experience/experimentation, was that the modes' frequencies, per month, in most cases, exhibited monotonically increasing and decreasing magnitudes. An example will be also provided with the individual calendar month unit changed to a half-month, doubling the number of intervals.

The results, cluster by cluster, are depicted graphically on single-chart layouts, the hourly mean vector wind directions as arrows, proportional in length to their magnitude (the highest magnitude vector is annotated), "constancy" statistics (mean vector wind speed/mean scalar wind speeds*100; a measure of vector wind persistency) as colored circles (ordered from dark blue to dark red), and month-by-month cluster relative frequencies, as a vertical bar-chart.

Since mean scalar wind magnitudes are not depicted on the charts, the constancy magnitude shadings can be utilized to estimate them, if desired, from the vector wind magnitude depictions. For example, constancy values that are say, 90 (or higher - a frequent result seen in clusters' output), would require a correction factor of 1/.90 or just 1.11 (or less - a mean vector wind speed annotation of 10 knots, for example, would be converted to a mean scalar wind speed of 11.1 knots). In such cases the vector wind info could be effectivley interpreted as a combined average wind direction AND scalar speed.

In other less ideal cases (with lower constancies), a results array might portray a diurnal pattern that had inherently lighter and more variable winds (for example, weak diurnal circulation regimes), and/or one with natural diurnal breaks (for example, land-breeze/sea-breeze transitions and vice-versa). The "average wind direction/wind speed" intepretation of the high constancy cases might not be as applicable here, but patterns of this nature, being an integral part of the given station's diurnal climatological "landscape" are, of course, no less important to present and describe.

3. RESULTS (SINGLE STATION ANALYSIS)

3.1 Point Mugu Naval Base (near Oxnard) - Lat: 34° 07' N; Lon: 119° 07' W; Elev: 13 ft.

Point Mugu (part of the Naval Base Ventura County Complex) is a Naval Air Station situated on the south-eastern edge of the Oxnard Plain in Ventura County, about 50 miles northwest of Los Angelea, and adjacent to the Santa Monica Mountains' western terminus (see Figure 1 map). As a coastal station bordering on the Pacific, it is subject to local and synoptic wind patterns associated with land breeze/sea breezes, downslope wind episodes ("Santa Anas) ", frontal passages, Paciific storms, and Catalina Eddies. A total of 12683 intact midnight-to-midnight days' wind observations were available to produce the six clusters. Figures 2 to 7 depict the idealized patterns, in rank order of frequency. They consist essentially of three flavors of land-breeze/sea-breeze, Santa Anas, south-southeasterlies, and strong westerlies episodes associated with strong sea-breezes/frontal passages.

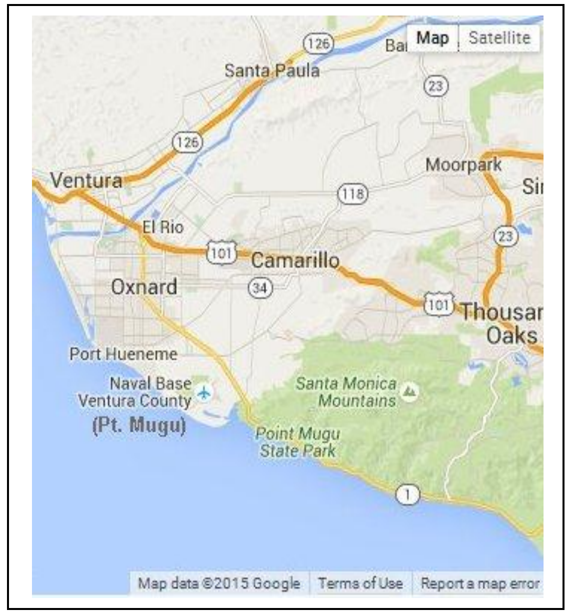

Figure 1 – Google Map of Pt. Mugu Naval Base And Surrounding Area

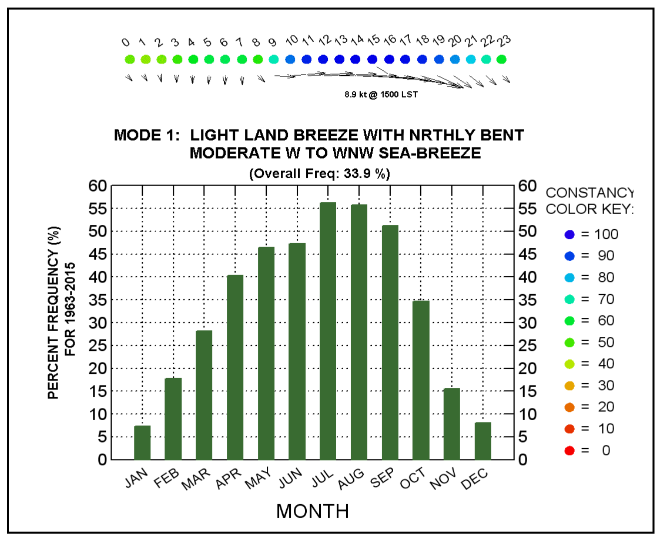

Figure 2. Pt. Mugu Mode #1 Graphical Depiction

3.1.1 *- Mode 1 – "Light Land Breeze with Nrthly Bent – Moderate W to WNW Sea Breeze " - (33.9% overall incidence)*

Figure 2 shows the most prominent Pt. Mugu idealized diurnal wind regime, a land-breeze/sea-breeze pattern that comprises 33.9% of the observations overall. It displays a monotonically increasing and decreasing progression of relative frequencies month-to-month, from about 7 % for January to more than 55 % for July and August. The 55 % figure indicates that 55 % of all July and August days had patterns that conformed statistically to this particular mode. Mean vector winds are moderately persistent (green constancy shadings) and light north-northwesterly to northerly through about 0900 LST, at which time a westerly sea-breeze takes over. This persists at high (blue constancy) levels from 1000 LST through about 1900 LST. The highest magnitude mean vector, as annotated in the chart, is a westerly one at 8.9 knots for 1500 LST. The constancy for this case is a very high 96, and using the "average wind direction/wind speed" interpretation applicable for high constancy values such as this one, the 1500 LST mean vector is effectively in mean scalar terms an average westerly wind at about 9 knots.

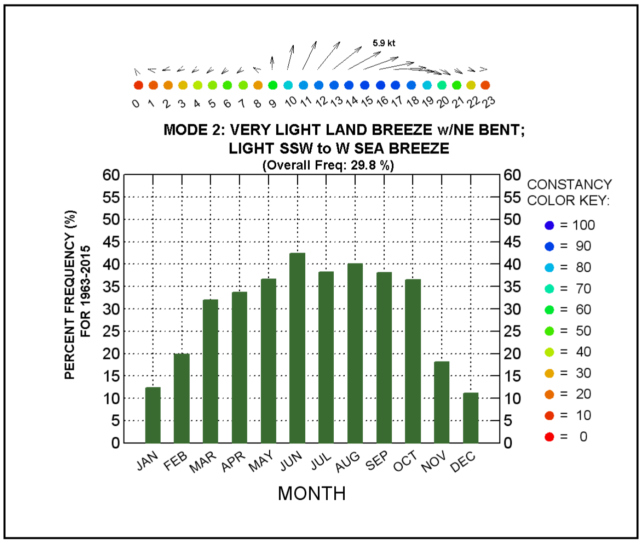

Figure 3. Pt. Mugu Mode #2 Graphical Depiction

3.1.2 - Mode 2 – *"Very Light Land Breeze with NE Bent – Moderate W to WNW Sea - Breeze"*

Figure 3 presents the second most prominent Pt. Mugu diurnal wind mode, another land-breeze/sea-breeze pattern that comprises 29.8% of the observations overall. It exhibits none of the peakedness of the Mode 1 bar chart (of which the frequencies for July-September were each greater than 50 %), presenting a more uniform distribution of magnitudes, especially from March through October, none higher than 42% (June). With the exception of June-July and July-August, the bars again show monotonically increasing and decreasing frequencies. The mean vectors display slight northeasterly orientations at very low constancies (orange to light green shadings) from 0300 to 0700 LST, shifting to light southerly at 0900 LST, then becoming stronger over the succeeding hours accompanied by a clockwise turning also accompanied with increasingly blue constancy shadings. Highest mean vector magnitude, a modest 5.9 knots (west-southwesterly) is attained at 1400 LST; the actual constancy statistic (89) thus indicates that in mean scalar terms this is a mean west-southwesterly at 6.6 knots. In contrast to Mode 1, this regime likely reflects a weaker regional west-to-east circulation pattern, local southerly sea-breeze influences predominating early (the Pt.Mugu coastline faces south) transitioning to the more regionally influenced and typical westerly ones as the day progresses. At 1700 LST, the mean vector orientation is westerly at 5.1 knots (constancy: 89).

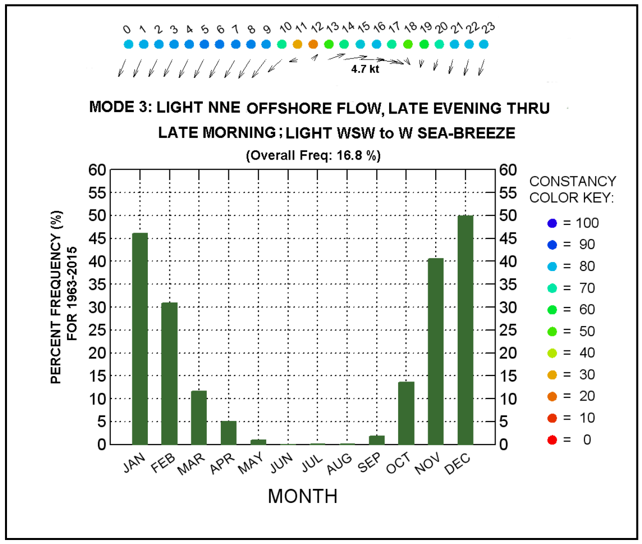

Figure 4. Pt. Mugu Mode #3 Graphical Depiction

3.1.3 - Mode 3 – *"Light NNE Offshore Flow, Late Evening thru Late Morning; Light WSW to W Sea-Breeze*

Figure 4 displays the third ranking mode, yet another land-breeze/sea-breeze pattern – the Pt. Mugu low-sun, cold- season version. Overall frequency is 16.8%, but nearly all are concentrated for the four months November thru February. December's frequency is 50%, January's 46%; the month-to-month progression of percentage statistics is completely monotonic.

The pattern is distinguished by light (3 to 4 knot) north-northeasterly mean vectors prevailing for most of the day, from 0000 LST to 0900 LST and 1900 LST to 0000 LST, at constancies typically in the high 70's to mid 80's. A brief sea-breeze period sets in over the early to late afternoon, the maximum magnitude vector, a westerly oriented one at 4.7 knots occurring at 1500 LST. Associated with a 74 constancy, this translates into a 6.3 knot mean scalar speed.

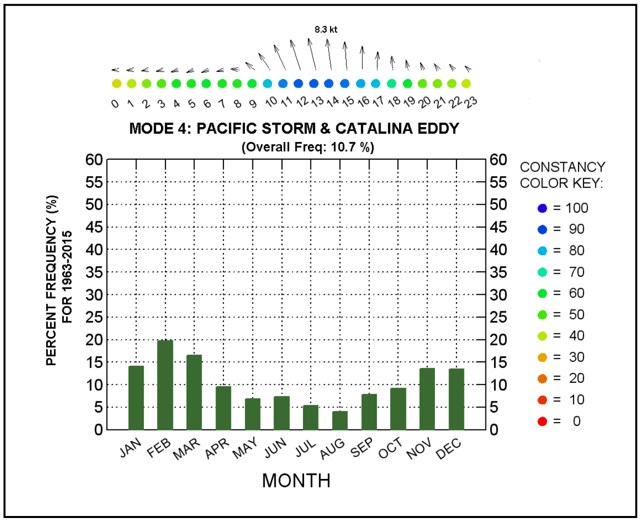

Figure 5. Pt. Mugu Mode #4 Graphical Depiction

3.1.4 - Mode 4 – *"Pacific Storm & Catalina Eddy"*

Fourth ranking mode is the "Pacific Storm & Catalina Eddy" pattern (Figure 5 – overall frequency 10.7%). This cluster's membership is likely a mixture of cases whose wind character is influenced by two separate weather-making entities which can produce south-easterly winds: Pacific storms and Catalina Eddies.

Pacific storms at Pt. Mugu typically occur over November to April with no known tendency for their southeasterly winds, when they do occur as part of the larger event, to have a diurnal dependence.

The Catalina Eddy, as described in the NOAA National Weather Service Glossary: "… forms when upper level large-scale flow off Point Conception interacts with the complex topography of the Southern California coastline. As a result, a counter clockwise circulating low pressure area forms with its center in the vicinity of Catalina Island. This formation is accompanied by a southerly shift in coastal winds, a rapid increase in the depth of the marine layer, and a thickening of the coastal stratus. Predominately these eddies occur between April and September with a peak in June". The Eddies can sometimes be a morning feature, with their peak winds occurring in the late morning or early afternoon, being dissipated in the hours following by the more normal westerly or northwesterly onshore flow patterns.

The mean vectors in Figure 5 show high constancy (upper 80's) southeasterly to southerly orientations at relatively high magnitudes over the late morning to mid-afternoon hours, the peak magnitude, an 8.3 knot south-southeasterly occurring at 1300 LST; also there are 5 to 10 % cluster frequencies over the May through October, months in which Pacific storms are rare or have never occurred.

So while Catalina Eddy occurrences are certainly represented in the frequency statistics of Figure 5, it would have been desirable to disentangle the two features by some means; perhaps doing a cluster analysis on the Pacific Storm & Catalina Eddy cluster itself. To this end, Figures 5a and 5b show results of this second order or "nested" clustering procedure of Mode 4, the number of clusters fixed at two.

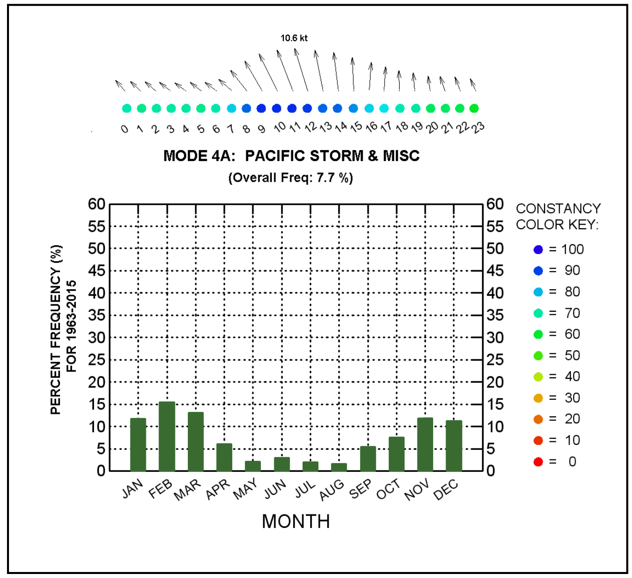

Figure 5a. Pt. Mugu #4a Graphical Depiction

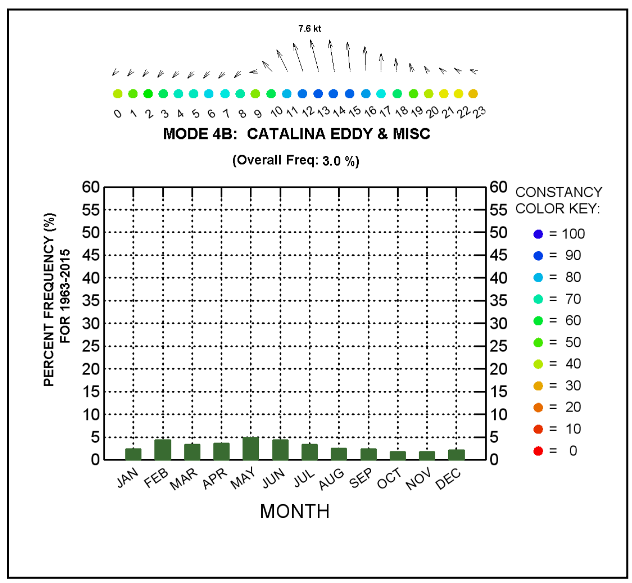

Figure 5b. Pt. Mugu #4b Graphical Depiction

The nested clustering operation produced two subclusters which in relative month-to-month terms seem to conform in rough fashion to the known anectodal relative frequencies of "Pacific Storms" and "Catalina Eddies" throughout the year.

Figure 5a, depicting the "Pacific Storms" pattern (7.7 % overall frequency) displays southeasterly to south-oriented vectors for virtually all hours of the day. The strongest vector (10.8 knots) observed at Noon PST.

The positioning of the maximum magnitude vectors between the hours of 0800 to1600 LST may be more an artifact of the clustering algorithm rather than indication that the highest Pacific Storm winds tend to gravitate towards the middle if the day; this may be an *average* hourly interval encompassing the highest winds.

Relative frequencies range between10 to 15 percent for the months November to March, conforming very well in incidence to those of Southern California rain-day events. Mid-summer (June-August) frequencies, while close to zero, seem to be a bit higher than expected, so perhaps cases reflecting additional very obscure patterns which have winds of a southeast to south orientation are included, hence the “Pacific Storm & Misc” title.

In contrast, Figure 5b shows light, offshore, northeasterly vectors from midnight to 0800 LST, followed by a shift to southeasterly to south orientations which cover most of the remaining hours of the day at varying magnitudes - this reflects to some degree the signature of Catalina Eddies. Highest magnitude is a 7.6 knot south-southeasterly at 1300 LST, 30 % less than the maximum showed in Figure 5a. Overall frequency is just 3.0%, with a much more uniform pattern displayed month-to-month, from about 2 to 5 percent, the relative maxima apparent for May and June.

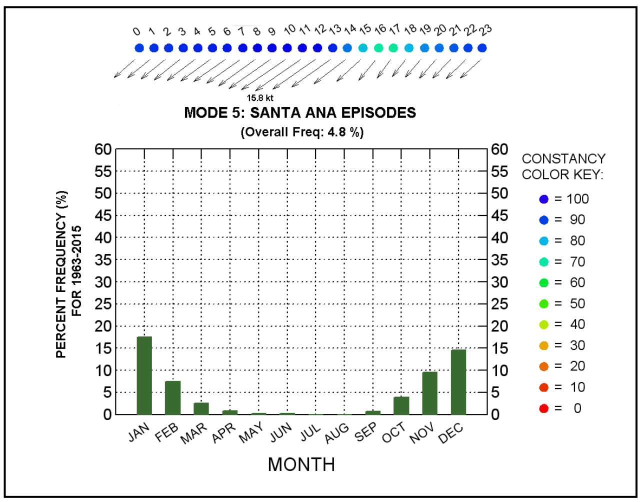

Figure 6. Pt. Mugu Mode #5 Graphical Depiction

3.1.5 - Mode 6 – *“Santa Ana Episodes”*

Ranking fifth is the “Santa Ana Episodes” pattern (Figure 5: overall incidence: 4.8 %), depicting the nature and monthly (monotonic) relative frequencies of these strong and usually very dry offshore downslope wind events that occur during the colder months of the year. Maximum frequencies are for January (17 %) and December (15 %). The mean vectors are northeasterly throughout the day at almost exclusively high (dark blue) constancy levels, the maximum magnitude,15.8 knots, noted for 1100 LST; this has a 97 constancy value. In the late afternoon (especially 1600 LST and 1700 LST), the constancies drop to around 70, probably reflecting some cases in which a brief sea-breeze period takes over, these being not frequent enough, however, to alter the mean vector orientation significantly. In the hours following, the offshore flow predominance resumes.

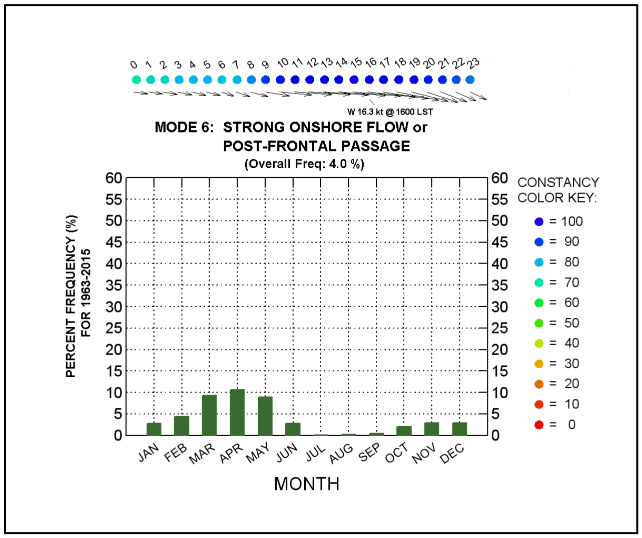

Figure 7. Pt. Mugu Mode #6 Graphical Depiction

3.1.6. - Mode 6 – *“Strong Onshore Flow or Post-Frontal Passage”*

The last cluster created was the “Strong Onshore Flow or Post-Frontal Passage” pattern (overall incidence: 4.0 %), grouping those relatively few cases which exhibited strong west to west-northwesterly wind events, especially for the afternoons. As Figure 7 indicates, such events are concentrated over March to May (frequencies around 10 %, each). This probably reflects the increasing west to east temperature and pressure gradients during Spring , the inland areas heating up rapidly with the increasing sun angles, the air over the oceans still relatively cool. Highest magnitude vector is a westerly 16.8 knot for 1600 LST; this has a 97 constancy.

4. RESULTS (TWO-STATION ANALYSIS)

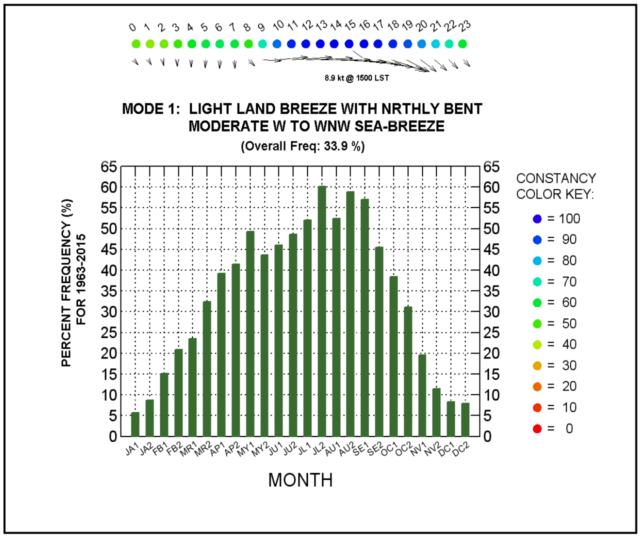

Figure 8. Pt. Mugu Mode #1 Bar-Graph Depiction in Half- Month Interval Terms

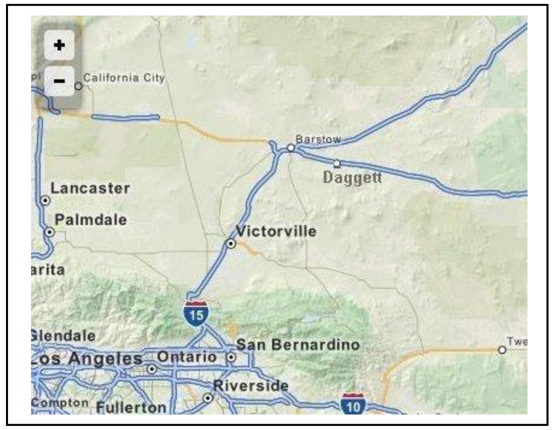

Figure 9 – Map of Victorville and Daggett and Surrounding Area (From City-Data.com)

3.1.7. Mode 1's Relative Frequencies at Half-Month Intervals

Two appealing visual properties of the graphs are the smooth hour to hour transitionings of the mean vectors and the already mentioned almost exclusively monotonic progressions of the monthly relative frequencies. As an experiment, Mode #1's cases are subgrouped into half-month classes, doubling the number of bars. Sample sizes of the half-month groups still range from 498 to 585. Results in Figure 8 show that the complete period-to-period monotonicity has gone away in some instances, but most of it remains. Maximum cluster relative frequency is now 60.2 %, for the second half of July group ("JL2").

In summary, with the exception of cluster 4 ("Pacific Storm and Catalina Eddy"), which required an additional "nested" cluster treatment, the K-means/V-fold analysis clearly resolved and delineated five mean vector diurnal modes for Pt. Mugu, this reflected by: 1) the relative ease of their physical interpretations and 2) the smooth monotonic progressions of their relative frequencies across the year.

An interesting but not altogether surprising takeaway, given Mugu's location, is that the highest three ranking patterns (all land-breeze/sea breeze modes), made up nearly 81 % of all the cases.

4.1 *- Daggett and Victorville Combined Analysis*

Figure 9 above is a map of Daggett, Victorville, and vicinity, two stations in the California High Desert northeast of Los Angeles. These were selected for their lengthy overlapping histories (covering 1951-2014), and their different geographical settings compared to Pt. Mugu.

Victorville (elev. 2875 ft), formerly George Air Force Base, and now the Southern California Logistics Airport (SCLA), is located about 25 miles northeast and downslope of the 3777 ft. elevation Cajon Pass. It exhibits relatively frequent southerly component winds, likely relating to its proximity to the San Gabriel Mountains to the west, the San Bernardino Mountains to the southwest, and the Cajon Pass itself.

Daggett Airport (elev. 1930 ft) is a public facility, located five miles east of Daggett proper, and 14 miles east of Barstow, the latter (elev. 2175 ft) 40 miles further northeast of Victorville. Somewhat farther removed from the topographical factors than influence Victorville, it shows a greater tendency for westerly winds.

Some 8716 total cases with complete simultaneous midnight-to-midnight wind observations were available for a 96-dimensional clustering analysis, and Figures 10 through 15 show the six patterns' results in descending order of relative importance.

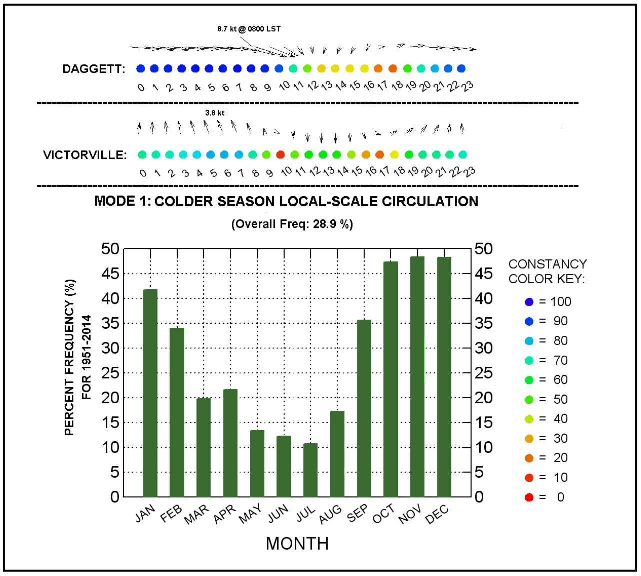

Figure 10 – Daggett/Victorville Mode #1 Graphlcal Depiction

4.1.1 - Mode 1 – *"Colder Season Local-Scale Circulation" - (28.9% overall incidence)*

Figure 10 displays the highest ranking mode for Daggett/Victorville, a cold season predominant pattern (percent frequencies all above 40 % for the months October through January) reflecting likely a weak regional circulation with local topographical influences, especially at night, prevailing in importance. Overall incidence is 28.9 %,

Daggett displays its maximum magnitude (westerly) vectors during the nocturnal and early morning hours, the absolute maximum, an 8.7 knot west-northwesterly combined with a 94 constancy, observed for 0800 LST; this could be a drainage wind. Mean vectors between 2200 LST and 1000 LST are all at high blue constancy levels with mean scalar speeds typically in the 8-9 knot range. In marked contrast, vector winds during the afternoon are light and variable, showing a very slight northeasterly bent; at these hours mean scalar speeds are only in the 5-6 knot range.

Victorviile has light relatively persistent south-easterlies (constancies mostly in the 70's) for the nocturnal hours (drainage wind from the Cajon Pass area?). At 1000 LST a shift takes place (red constancy shadings), to light northerlies, these persisting until 1600-1700 LST when they shift back to southwesterly, then southerly.

Given this is a predominant Fall/Winter regime, perhaps it reflects scenarios with a weak high pressure system situated over the Great Basin accompanied by slight pressure gradients. This results in hints of an offshore component during afternoons , but these are too weak to hold off at the surface the local topographical drainage winds(?) that reassert themselves nocturnally.

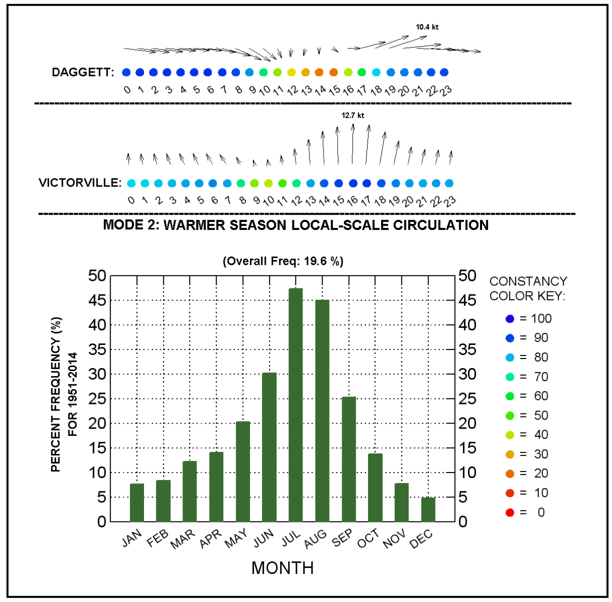

Figure 11 – Daggett/Victorville Mode #2 Graphlcal Depiction

4.1.2 - Mode 2 – *"Warmer Season Local-Scale Circulation" - (19.6% overall incidence)*

Second in importance, the "Warmer Season Local-Scale Circulation" pattern (Figure 11), is a sort of opposite-season "local-scale" counterpart to Figure 10, contrasts in features, however, shown between the two.

Clearly a mid-summer phenomenon, relative frequencies reach 47% and 45%, respectively for July and August. Overall frequency is 19.6 %.

At Daggett, the nocturnal westerlies are present again from about 2200 LST to 0700 LST (drainage wind?) at blue constancy shadings and mean scalar speeds in the 8-9 kt range. Over 0800-1100 LST they briefly transition to more northwesterly, then through 1500 LST being quite variable (brown to yellow low constancy colors); mean scalar speeds over these high-sun hours are only 6-7 kt. In the late afternoon and after, the winds become stronger and more persistent direction-wise, an early evening (1900 LST) southwesterly maximum of 10.4 kt and 84 constancy noted (corrected to a12.5 kt mean scalar speed). Thereafter, the magnitudes lessen and become more westerly.

At Victorville, light southeasterlies with high 70's to low 80's (light blue to blue) constancy values (drainage winds?) prevail from midnight to 0700 LST. Then, around 0900 LST and, interestingly, several hours earlier than Daggett, a transition interval ensues (light green to yellow constancy shadings through 1100 LST with progressively stronger hour-to-hour southerlies developing. Maximum vector magnitude is noted at 1600 LST, a southerly at 12.7 kts with an associated constancy of 91 ("adjusted" to a 14.3 kt mean scalar speed).

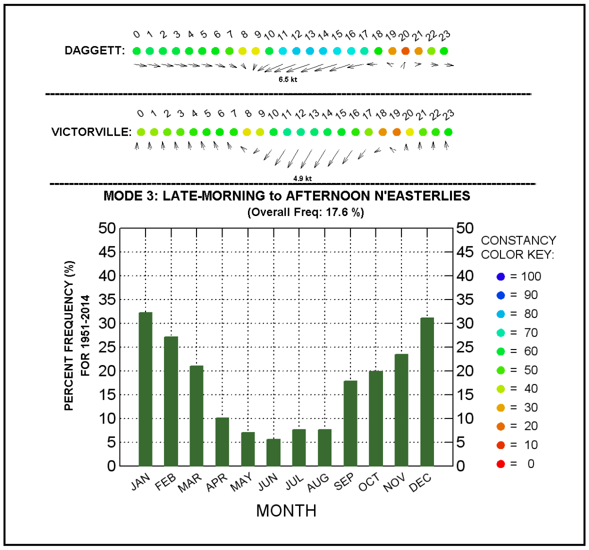

Figure 12 – Daggett/Victorville Mode #3 Graphlcal Depiction

4.1.3 - Mode 3 – *"Late-Morning to Afternoon N'Easterlies"- (17.8% overall incidence)*

Third in rank is a lesser cold-season type mode, the "Late-Morning to Afternoon N'Easterlies" pattern (Figure 12). Having an overall frequency of 17.8%, this is marked by the light east northeasterly to northeasterly vectors displayed at both stations over the late morning to late afternoon hours, and the relatively low, in general, vector magnitudes and low constancies hour-to-hour. Average figure at Daggett is only 57, that for Victorville 49.

Daggett displays its "usual" light westerly mean vectors from 2200 LST thru 0700 LST, but following a two-hour transition period (yellow shaded low constancies for 0800 and 0900 LST), an hour-to-hour run of east-northeasterlies sets in, prevailing through roughly 1800 LST. Following another three-hour transition period covering 1900 LST to 2100 LST (orange low constancy shadings), the light westerly nocturnal pattern resumes. Maximum vector magnitude is a 6.5 kt east-northeasterly at 1300 LST, the associated constancy 78 with the scalar speed adjustment, 8.3 kt.

Victorville also displays its "usual" light southerly vectors over 2200 LST to 0700 LST, and after an 0800 LST to 0900 LST transition period (yellow low constancy shadings), identical to Daggett's, an hourly succession of northeasterlies prevails into the late afternoon. Like Daggett, a three-hour transition period (orange and yellow low constancy shadings) then ensues, but somewhat curiously, this commences one hour earlier than Daggett, at 1800 LST. Maximum magnitude northeasterly vector for Victorville is 4.8 kt at 1300 LST, associated with a constancy of 64 and a adjusted mean scalar speed of 7.7 kt.

The occurrence of these northeasterly quadrant orientations at both stations suggests that this particular mode may be associated, on a larger regional basis, with Great Basin Highs that are, in contrast with Mode 1, possesed of stronger pressure gradients. Some of these individual cases, especially in the Fall and Winter months, may be associated with Santa Ana events downstream in the lower elevation coastal regions to the West.

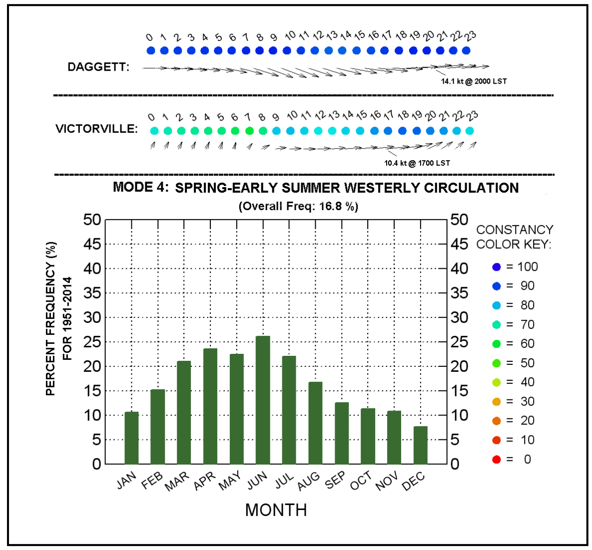

Figure 13 – Daggett/Victorville Mode #4 Graphlcal Depiction

4.1.4 - Mode 4 – *"Spring Early-Summer Westerly Circulation" - (16.8% overall incidence)*

Ranking fourth (Figure 13) is the "Spring-Early Summer Westerly Circulation" pattern (16.8% overall incidence). This might reflect diurnal wind patterns expressed in the wake of frontal passages, not necessarily immediately following but some period of time after. Maximum frequencies occurr for the months April-June.

The Daggett vectors are at blue constancy levels (range; 87 to 94) for all hours of the day, the vector maximum magnitude noted for 2000 LST, a westerly at 14.1 knots with a constancy of 94 (adjusts to a 15 knot mean scalar speed).

Likely shielded somewhat from the onshore flow by the local mountainous topography, the Victorville pattern, in contrast, shows light southwesterlies from midnight to 0800 LST (constancies in the 60's and 70's and mean scalar speeds in the lower 5 kt range). From 0900 LST on the vectors are westerly, increasing in magnitude through 1700 LST when the maximum 10.4 kt at a constancy 85 is recorded (adjusts to a 12.3 kt mean scalar speed). From 1900 LST on they are back to southwesterly. Interestingly, with the stronger synoptic scale entity in play here, unlike the previous three patterns, there are no mean vector 'transition" hours (those with very low constancy yellow, orange, or red shadings) seen in the graph, these local diurnal effects being overwhelmed.

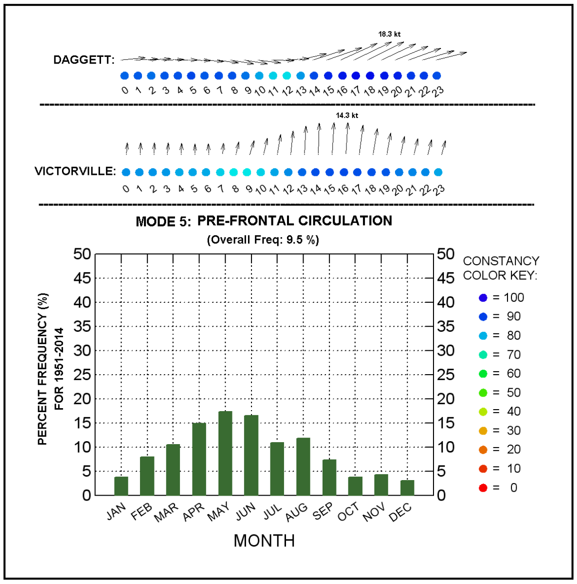

Figure 14 – Daggett/Victorville Mode #5 Graphlcal Depiction

4.1.5 - Mode 5 – *"Pre-Frontal Circulation Circulation" - (9.5% overall incidence)*

In fifth place (Figure 14) is the "Pre-Frontal Circulation" pattern (overall incidence: 9.5 %). Once again, the relative frequencies in this case are concentrated in April-June. This reflects the strong winds that can occur in the High desert region in advance of a frontal passage.

In Daggett the mean vectors through 0900 LST are westerly at constancies in the 80's and mean scalar speeds in the 9-11 knot range. Over the 1000 LST to 1200 LST interval they shift (note the lighter blue constancies) to progressively more southwesterly with higher magnitudes; maximum magnitude vector is seen at 1700 LST, a southwesterly at 18.3 knots at a constancy of 95 (adjusts to a 19.2 mean scalar speed). At 2300 LST they are still southwesterly at 14 knots and at an 88 constancy.

The Victorville vectors are mostly southerly throughout the day at varying blue constancy levels (like Daggett). There is a bit of a transition period over 0700-0900 LST (note the lighter blue shading), the vectors becoming temporarily more south-southwesterly, but therafter returning to southerly, exhibiting a maximum mean speed of 14.3 kt at 1600 LST (southerly at an 88 constancy, 16.2 kt mean scalar speed adjustment).

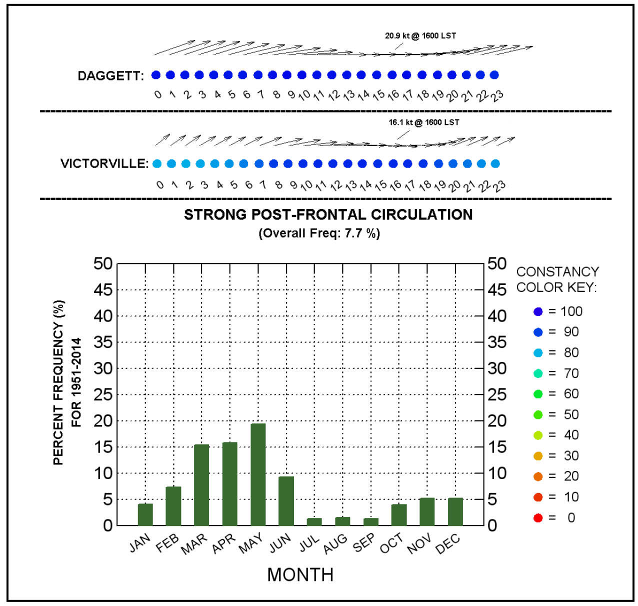

Figure 15 – Daggett/Victorville Mode #6 Graphlcal Depiction

4.1.6 - Mode 6 – *"Strong Post-Frontal Circulation" - (7.7% overall incidence)*

Sixth mode is the "Strong Post-Frontal Circulation" pattern (Figure 15: overall frequency 7.7%). An analog to Point Mugu's Figure 7 pattern, strikingly similar in bar-graph configuration, this reflects those comparatively infrequent very strong frontal passage scenarios that are confined mostly to the spring months, March-May. The May percentage (20 %), in particular, indicates some 6 event occurrences per month on average.

The Daggett and Victorville vectors for this mode are more alike in orientation and constancy level than any of the other previously discussed five, although Daggett's are generally more west-southwesterly and higher in magnitude. At each station they become almost due westerly in the afternoons.

Daggett's constancy levels are at high dark-blue levels throughout the day; seven mean vectors having magnitudes in the 20 kt or higher catefgory. Absolute maximum magnitude, recorded at 1600 LST, is a 20.9 kt westerly with a constancy of 94 (adjusted mean scalar speed: 22.2 kt).

Victorville's constancies are not as consistently high as Daggett's, somewhat lower in the 80's over 1900 LST thru 0700 LST, but at dark-blue levels for the other hours. Maximum magnitude vector is a 16.1 kt westerly at 1600 LST and constancy 92 (adjusted mean scalar speed: 17.4 kt)

In summary, the Daggett/Victorville analysis resolved a series of six modes that essentially described three relatively weak circulation patterns (the top ranking three – making up 66.3 % of the cases), and three lesser ones that characterized synoptic pattern episodes that were associated with significant winds.

Like Pt. Mugu, the relative month-to-month frequencies showed mostly monotonic variability.

Interpretations of the patterns was necessarily more complicated and subjective since two stations were being considered, but in most cases it seemed to be a relatively straightforward exercise, attempts made to explain the most prominent and significant features. It is of course, certainly possible that lesser modes, with relative frequencies concentrated over a few contiguous months but with physically meaningful explanations await resolution.

5. CONCLUSION

Utlizing K-Means Clustering Analysis combined with an efficiency improving software add-on called the V-Fold Cross Validation Algorithm, the foregoing analyzed, as two separate test excercises, decomposed u and v hourly wind components for a single station (Pt. Mugu, CA), and two stations combined (simultaneous u and v wind observations for Daggett and Victorville, CA), the objective being to create useful idealized diurnal wind pattern characterizations in the form of mean vector wind data arrays that encompassed, hour-to-hour, the midnight-to-midnight period. The single station results confirmed and reinforced those from previous such analyses, the two-station analysis for Daggett and Victorville, the first of its kind, producing similarly favorable conclusions.

The single and double station results were both presented in a newly introduced graphical template (one chart per idealized pattern or mode) and two feature portrayals on the graphs that seemed to validate the clustering methodology were 1.) the smooth progression of the climatological mean vector depictions hour-by-hour, permitting plausible physical interpretations and 2.) the smooth monotonic progressions of the various patterns' climatological frequencies across the year.

A takeaway of the two station analysis, with wind data available for number of other lengthy-history stations in the High Desert area, like Lancaster Palmdale, and Edwards Air Force Base, one could expand the cluster treatment to five stations, providing a more complete regional picture of the different flavors of simultaneous diurnal wind variability; a downside, though, might be increased interpretative complexities.

6. REFERENCES

Fisk, C.J., 2012: "Identification of Diurnal Wind Pattern Modes, Optimally Numbered, for Four Coastal Southern California Stations Utilizing the V-Fold Cross-Validation Algorithm Applied to K-Means Clustering Analysis – 10th Conference on Artificial Intelligence Applications to Environmental Science, American Meteorological Society, 22-26 January 2012, New Orleans https://ams.confex.com/ams/92Annual/webprogram/Paper192739.html

Nisbet, R., Elder, J., and Miner, G., 2009: Handbook of Statistical Analysis & Data Mining Applications Elsevier, 824 pp.