Studying Important Weather Factors in the 2021 Texas Power Outage Crisis

Mason Chen[1]
[1]Stanford OHS, 415 Broadway, Redwood City, CA 94063

Abstract
This paper studied 2021 Feb. Texas Power Outage due to the Winter Storm Uri. Houston was hit much harder than Dallas even both has the same ERCOT power system. Cluster Sampling method was adopted and collected 2012-2021 Feb. Houston daily weather data. To improve the modeling capability, four primary weather factors were selected based on the cluster variables algorithm. Several Statistical JMP Platforms such as Outlier, SPC Control Chart, PCA, Heat Map, Score Plot were conducted to analyze the Houston weather data. Average Temperature and Dew Point were found to be the most contributing weather factors to 2021 Feb. Houston Power Outage Crisis. The 2021 Feb. outage happening probability is lower than 0.5% (once in more than 7 years).

Key Words: Texas Power Crisis, Dew Point, Statistics, SPC, Clustering

1. Introduction

Texas State's failure to winterize power equipment left natural gas facilities and wind turbines vulnerable to the extreme winter storms in February 2021. Though, recommendations already made by the North American Electricity Reliability Corporation as a result of the 2011 blackouts were ignored due to cost considerations. Many oil producers were willing to take the risk of possible "freeze-ups". The Electric Reliability Council of Texas (ERCOT) decided to impose rolling blackouts on Texas residents, leaving many in crisis conditions, resulting in 2021 Feb. Power and Water outage crisis never in history.

1.1 Texas Crisis Timeline

"Winter Storm Uri" hits Texas, already suffering from freezing temperatures on Feb.13, 2021. Electricity demand hit 69,150 MW on Feb.14, 3,200 higher than the record set in 2018. Feb.15, ERCOT begins rotating power outages, 4.4 million people are left without power. ERCOT ends emergency conditions and ceases rolling blackouts on Feb.19, 2021.

1.2 Scope and Objective

The main objective of this paper is to find out which weather factors such as humidity, temperature, and dew point had the greatest impact on the Houston outage situation. Utilize Quantile Range Outliers, Principal-Component Statistical Process Control Chart, and Cluster Variables statistical methods to determine which markers differed the most from past years. Compare the 2021 climate to the past decade using the Heat Map and the Score Plot. Connect the outlier analyses results to environmental science to ultimately determine the most representative indicator(s) of similar crises in the future.

1.3 Crisis in Houston more than in Dallas

What weather factor had the greatest contribution? Houston was hit much harder, during 2021 Feb., by the crisis than Dallas (but both used the same power system). Houston has higher temperature and higher dew point as shown in Figure 1. Is the dew point the most important weather indicator of "freeze-up" risks?

Houston

Temperature (° F)	Max	Average	Min
Max Temperature	82	61.68	27
Avg Temperature	74.82	53.32	21.67
Min Temperature	71	45.68	15
Dew Point (° F)	Max	Average	Min
Dew Point	69	43.88	3

Dallas

Temperature (° F)	Max	Average	Min
Max Temperature	79	51.61	14
Avg Temperature	65.31	42.45	9.47
Min Temperature	56	34.11	3
Dew Point (° F)	Max	Average	Min
Dew Point	66	30.57	-5

Fig.1 Temperature and Dew Point: Houston vs. Dallas

1.4 Weather Science

Air temperature (commonly known as just temperature) is the measure of the average kinetic energy of air molecules (condensation depends on air temperature). Relative humidity (commonly known as just humidity) expresses how much energy available has been used for evaporation. Dew point temperature (commonly known as just dew point) is the temperature at which condensation first begins (measure of moisture). Humidity depends on both temperature and dew point (which are both independent variables) as the greater the difference between air temperature and dew point, the lower the relative humidity. Dew point is independent of atmospheric pressure, but temperature and humidity are directly related to pressure. Could the weather science explain the crisis severity between Houston and Dallas in 20121 Feb.?

2. Data Collection

To investigate the weather factors contributed to the 2021 Feb., Texas Power crisis. A Cluster sampling method was adopted by collecting the daily Houston weather statistics in the month of February from 2012 to 2021 [1]. The other months were not considered in this paper to improve the data integrity. The total Sample size = 283 (28 days *10 years + 3 days from leap years) should be sufficient to draw statistical decision at high confidence level. Weather parameters that were collected were temperature, dew point, humidity, wind speed, pressure, and precipitation of Houston City as shown in Figure 2.

Time	Temperature (° F)			Dew Point (° F)			Humidity (%)			Wind Speed (mph)			Pressure (Hg)		
Feb	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min
1	65	53.8	44	34	29.5	20	68	42.8	20	16	9.7	3	30.4	30.3	30.2
2	66	54.7	47	43	36.1	29	80	51.3	35	13	6.6	0	30.3	30.2	30.1
3	68	59.3	49	56	50.2	43	89	73.0	55	15	9.6	0	30.1	30.0	29.9
4	79	69.6	61	61	58.0	44	87	68.1	48	23	12.4	6	29.9	29.8	29.7
5	61	53.0	48	44	41.1	36	80	65.4	47	21	12.7	7	30.0	29.9	29.8
6	68	57.8	50	56	49.8	43	89	75.0	64	17	8.3	3	30.0	29.8	29.8
7	63	52.6	43	47	42.6	38	86	69.9	50	15	8.8	5	30.1	30.0	29.9
8	71	62.3	53	61	56.1	48	93	80.9	63	15	9.5	5	30.0	29.9	29.9
9	71	66.0	64	65	62.6	60	96	88.9	78	14	6.1	0	30.0	30.0	29.9
10	68	62.1	48	64	59.2	46	97	90.3	78	16	7.9	0	30.0	29.9	29.9
11	47	43.1	40	45	40.3	38	93	89.7	85	21	14.5	10	30.0	30.0	29.9
12	42	38.2	36	37	34.6	32	93	87.1	79	22	16.4	12	30.1	30.1	30.0
13	44	39.7	36	32	31.6	31	82	72.8	63	18	14.1	9	30.1	30.0	30.0

Fig.2 Data Collection of Houston Weather in Feb.

3. Statistical Analyses

In section 3: several JMP statistical platforms were utilized to further investigate what weather factors contributed to the 2021 Feb. Houston Power Outage Crisis.

3.1 Outlier Analysis

The Quantile Range Outliers analysis was conducted to detect the outliers across all weather parameters during 2012-2021 Feb. months. The outlier threshold setting was at 0.25 Tail and Q=1.5 as standard outlier detection method [2]. The Max Humidity and Perception were two parameters with most outliers detected during the 2012-2021 Feb. period. Though, this observation may not reflect what were really happening during 2021 Feb. since crisis period was very short in a few days.

Quantile Range Outliers

Outliers are values Q times the interquantile range past the lower and upper quantiles.

Tail Quantile 0.25

Q 1.5

Select columns and choose an action.

☐ Restrict search to integers

☐ Show only columns with outliers

Some quantiles were stretched to avoid a large group at the median.

Column	25% Quantile	75% Quantile	Low Threshold	High Threshold	Number of Outliers	Outliers (Count)
Temp-Max	62	76	41	97	6	27 36 38 39 40(2)
Temp-Avg	53	67.1	31.85	88.25	2	21.7 26.3
Temp-Min	43	59	19	83	2	15 18
Dew Point-Max	47	66	18.5	94.5	0	
Dew Point-Avg	39	61.3	5.55	94.75	0	
Dew Point- Min	32	55	-2.5	89.5	0	
Humidity%-Max	86	97	69.5	113.5	19	46 49 53 54(2) 55 56 59 63 64 65 66(2) 67(2) 68(2) 69(2)
Humidity%-Avg	64.5	87	30.75	120.75	1	29.9
Humidity%-Min	39	69	-6	114	0	
Wind Speed-Max	14	20	5	29	5	30(2) 31(2) 32
Wind Speed-Avg	6.9	11.4	0.15	18.15	1	18.4
Wind Speed-Min	0	5	-7.5	12.5	2	13(2)
Pressure-Max	30	30.3	29.55	30.75	0	
Pressure-Avg	29.9	30.2	29.45	30.65	0	
Pressure-Min	29.8	30.1	29.35	30.55	0	
Preception (in)	0	0.06	-0.09	0.15	21	0.16 0.19 0.22 0.23 0.26(2) 0.27 0.41(2) 0.43 0.61 0.7 0.71 0.76 0.89 0.92 0.93 1.13 1.41 1.48 1.82

Fig.2: Quantile Range Outliers Analysis

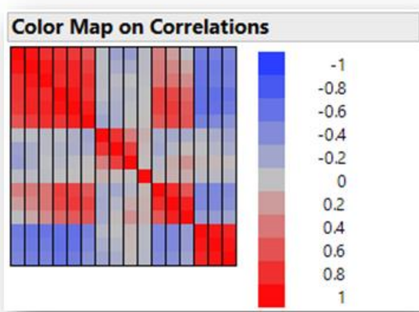
JMP Platform could colorize the detected outliers in the original raw data table as shown in Figure 3. Temp Max was the most sensitive weather factor which could match well with Power crisis period. Humidity Max was another factor to be considered.

Year	Feb	Temp-Max	Temp-Avg	Temp-Min	Dew Point-Max	Dew Point-Avg	Dew Point- Min	Humidity%-Max	Humidity%-Avg	Humidity%-Min	Wind Speed-Max
2021	2	66	54.7	47	43	36.1	29	80	51.3	35	13
2021	3	68	59.3	49	56	50.2	43	89	73	55	15
2021	4	79	69.6	61	61	58	44	87	68.1	48	23
2021	5	61	53	48	44	41.1	36	80	65.4	47	21
2021	6	68	57.8	50	56	49.8	43	89	75	64	17
2021	7	63	52.6	43	47	42.6	38	86	69.9	50	15
2021	8	71	62.3	53	61	56.1	48	93	80.9	63	15
2021	9	71	66	64	65	62.6	60	96	88.9	78	14
2021	10	68	62.1	48	64	59.2	46	97	90.3	78	16
2021	11	47	43.1	40	45	40.3	38	93	89.7	85	21
2021	12	42	38.2	36	37	34.6	32	93	87.1	79	22
2021	13	44	39.7	36	32	31.6	31	82	72.8	63	18
2021	14	38	34.1	28	33	28.8	24	86	80.9	70	22
2021	15	27	21.7	18	23	14.1	7	88	73.3	58	25
2021	16	36	26.3	15	26	12.5	3	67	56.5	42	14
2021	17	39	36.4	33	37	32.6	30	93	86.4	76	22
2021	18	44	35.8	32	33	27.7	20	92	73.2	58	23
2021	19	51	38.5	26	28	22.1	19	78	54.2	32	14
2021	20	60	46.2	32	43	35.3	27	83	67.8	44	14
2021	21	72	59.1	46	54	49.5	42	93	72.8	44	20
2021	22	76	61.8	51	54	40.6	24	90	53.9	15	13
2021	23	79	61.6	45	53	45.5	37	89	62.3	23	17
2021	24	80	70.1	58	66	62.1	52	90	76.8	56	17
2021	25	72	69.3	66	67	65.1	64	93	86.5	78	15

Fig.3: Map February 2021 Quantile Range Outliers.

3.2 Cluster Variables Analysis

To further investigate the weather factors, modern Cluster Variables analysis [3] was conducted to group the similar factors in a separate cluster as shown in Figure 4 Color Map and Cluster List analysis. Four clusters were identified and each weather factor was assigned to one cluster. Each cluster has also identified the most representative weather factor. The 1st cluster is Temp; the 2nd cluster is Wind Speed; the third cluster is Humidity; the fourth cluster is pressure. Each cluster may indicate one independent weather mechanism to explain the association weather patterns during 2012-2021 Feb. months. This cluster analysis is a good analysis to brainstorm the weather science but not sufficient to draw any meaningful insights regarding the 2021 Feb. Houston Power Crisis.



Cluster Summary

Cluster	Number of Members	Most Representative Variable	Cluster Proportion of Variation Explained	Total Proportion of Variation Explained
1	6	Temp-Avg	0.851	0.319
4	3	Pressure-Avg	0.941	0.176
3	3	Humidity%-Avg	0.845	0.158
2	4	Wind Speed-Avg	0.566	0.141

Fig. 4: Cluster Variables Analysis

Cluster members were further listed in Figure 5. The most interesting observation is that Temp and Dew Point were assigned to the 1st cluster at a very high correlation. As discussed in Section 1.4 weather science, the Dew Point is related to both the Temp and Humidity. Also, in Section 1.3, Houston had much higher Temp and Dew Point than Dallas. This observation may indicate that Dew Point, in addition to temp, may be another weather factor which contributed to the 2021 Houston Power crisis.

Cluster Members

Cluster	Members	RSquare with Own Cluster	RSquare with Next Closest	1-RSquare Ratio
1	Temp-Avg	0.922	0.316	0.114
1	Dew Point-Avg	0.923	0.535	0.166
1	Temp-Min	0.88	0.294	0.17
1	Dew Point-Max	0.87	0.457	0.24
1	Temp-Max	0.771	0.255	0.307
1	Dew Point- Min	0.741	0.457	0.478
2	Wind Speed-Avg	0.936	0.022	0.065
2	Wind Speed-Max	0.697	0.031	0.313
2	Wind Speed-Min	0.616	0.012	0.388
2	**Precipitation**	0.014	0.009	0.995
3	Humidity%-Avg	0.975	0.298	0.036
3	Humidity%-Min	0.807	0.135	0.224
3	Humidity%-Max	0.754	0.325	0.365
4	Pressure-Avg	0.975	0.343	0.038
4	Pressure-Min	0.927	0.273	0.1
4	Pressure-Max	0.921	0.377	0.127

Fig.5 Four Clusters

3.3 Model Driven SPC Control Chart Analysis

To further investigate the weather factor pattern in time domain, a modern Principal Component Analysis (PCA) driven Statistical Process Control (SPC) chart platform was conducted to analyze the combined T2 Multivariate control chart.

3.3.1 Default PCA-Driven SPC Control Chart Analysis

A default setting (include all weather factors) of PCA-driven SPC method [4] was conducted as shown in Figure 6. The 2021 Power Outage was located around Sample 270. Though, the default setting cannot detect the 2021 weather situation very well. It may be necessary to remove certain weather factors (Noise) to improve the control chart detection power.

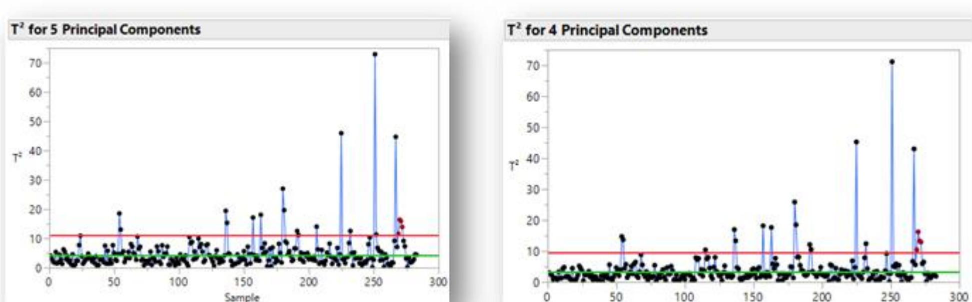


Fig.6 Default SPC Control Chart Analysis

3.3.2 Enhanced PCA-Driven SPC Control Chart Analysis

To improve the statistical signal-noise ratio, four primary weather factors were selected as shown in Figure 7 on the left chart. Average parameters are more representative than Minimum and Maximum which may make sense of the accumulated impact during the crisis period. The updated T2 control chart only needed 2 principal components (over 4 or 5 principal components in the default method) which has significantly simplified the complexity of the weather association patterns. The highest peak in the control chart was located at sample 270 which was exactly the worst date during the 2021 Mid. Feb. Houston Crisis. The four weather factors selected may be sufficient to explain well of Houston Power Crisis in 2021 Mid. Feb.

Cluster Members

Cluster	Members	RSquare with Own Cluster	RSquare with Next Closest	1-RSquare Ratio
1	Temp-Avg	0.922	0.316	0.114
1	Dew Point-Avg	0.923	0.535	0.166
1	Temp-Min	0.88	0.294	0.17
1	Dew Point-Max	0.87	0.457	0.24
1	Temp-Max	0.771	0.255	0.307
1	Dew Point- Min	0.741	0.457	0.478
2	Wind Speed-Avg	0.936	0.022	0.065
2	Wind Speed-Max	0.697	0.031	0.313
2	Wind Speed-Min	0.616	0.012	0.388
2	Preception (in)	0.014	0.009	0.995
3	Humidity%-Avg	0.975	0.298	0.036
3	Humidity%-Min	0.807	0.135	0.224
3	Humidity%-Max	0.754	0.325	0.365
4	Pressure-Avg	0.975	0.343	0.038
4	Pressure-Min	0.927	0.273	0.1
4	Pressure-Max	0.921	0.377	0.127

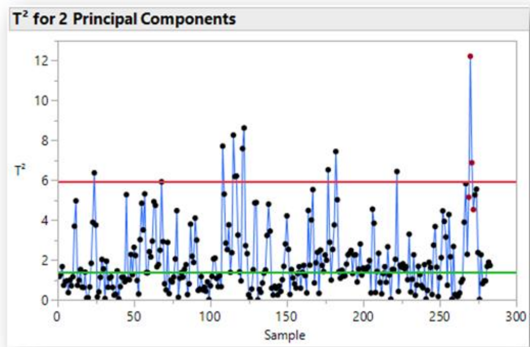


Fig.7 Enhanced Model Driven SPC Control Chart Analysis

3.3.3 T2 Contribution Proportion Analysis

To further investigate the control chart result, T2 Contribution Proportion Analysis was conducted to analyze which weather factor (among four selected) was the most contributed one during 2021 Feb. Houston Power crisis period as shown in Figure 8. Temperature and Dew Point factors are the main contributors to the weather situation of the 2021 outage crisis. The red columns shown were regarded to the unstable weather factors in Power outage period. This is another observation that Temperature and Dew Point were linked together to Power outage crisis.

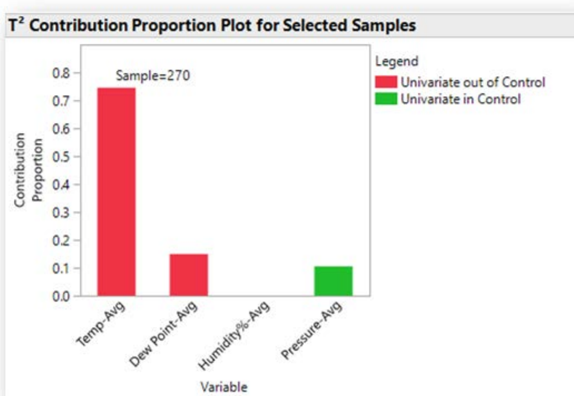


Fig.8 T2 Contribution Proportion Plot

3.4 Principal Component Analysis

In the enhanced PCA model driven Control Chart analysis, only two principal components needed to explain what happened to Houston Power outage mechanism. PCA analysis [5,6] was also conducted to draw more insights as shown in Figure 9. Principal Component 1 and 2 can explain more than 80% of the variation (Pareto principle). This finding has further supported the power of the enhance model by selecting four weather factors. First eigenvector is quite evenly distributed among the four parameters. Second eigenvector mainly considers temperature and humidity.

Eigenvalues

Score	Eigenvalue	Score Std Dev	Percent	20 40 60 80	Cum Percent
1	2.784	1.6684	69.593		69.593
2	0.701	0.8372	17.524		87.116
3	0.510	0.7142	12.753		99.869
4	0.005	0.0723	0.131		100.000

Eigenvectors

	Prin1	Prin2	Prin3	Prin4
Temp-Avg	0.50198	-0.55380	0.40120	0.52949
Dew Point-Avg	0.58275	0.04105	0.31481	-0.74807
Humidity%-Avg	0.44385	0.80166	0.02369	0.39972
Pressure-Avg	-0.45980	0.22128	0.85987	0.01582

Fig.8 Principal Component Analysis

3.5 Heat Map Analysis

Heat Map analysis was also conducted to visualize the pattern among four weather factors in time domain.

3.5.1 Heat Map Analysis across 2012-2021 Feb. Months

The heat map analysis was first conducted across the 2012-2021 as shown in Figure 9. Temperature and Humidity% are the main contributors across February 2012 to 2021. Though, we should look at the pattern during the 2021 Feb. outage crisis period since the heat map pattern may be quite differently.

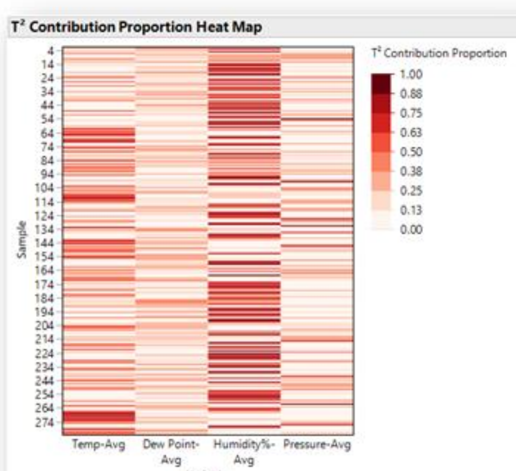


Fig.9 Heat Map of 2012-2021 Feb.

3.5.2 Heat Map Analysis in 2021 Feb.

The heat map analysis was duplicated during the 2021 Feb. crisis period as shown in Figure 10. Temperature and Dew Point are the main contributors for the 2021 outage crisis. The pattern during the 2021 Feb. crisis period is significantly different from the entire 2012-2021 period. This observation may indicate the historical 2012-2020 weather pattern may not predict the 2021 Feb. outlier crisis.

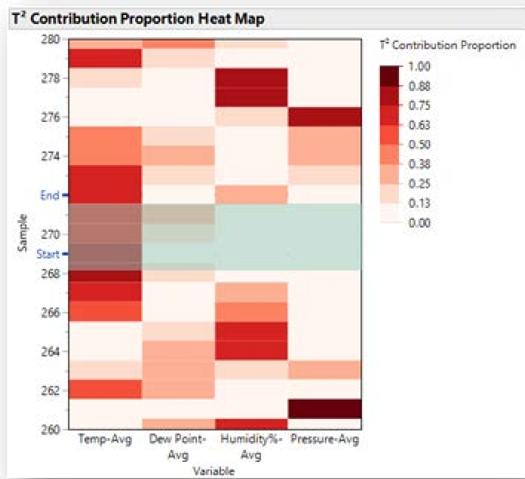


Fig.10 Heat Map during Feb. 2021 Outage Crisis

3.6 Score Plot Analysis

The Score Plot analysis can plot the probability at various Confidence Ellipse levels based on the first two Principal Components as shown in Figure 11. February 15th (the day with the worst weather situation) is outside the 99.5% confidence ellipse (occurs once every 7 years when disregarding other factors). This result is consistent with previous analysis that it's very difficult to predict the happening of the 2021 Feb. crisis based on 2012-2020 data. The four days with the worst weather situation lie opposite to the temperature and dew point eigenvectors which has indicated the extreme outliers of temp and dew point occurred during 2021 Feb. outage period.

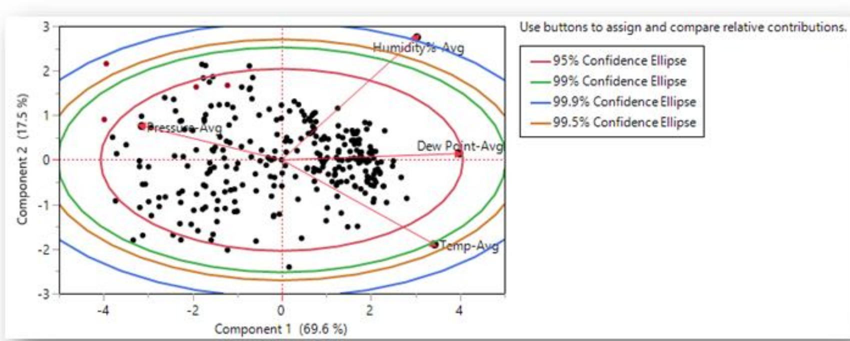
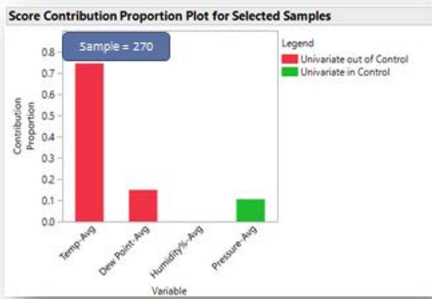


Fig.11 Score Plot Analysis

Next step is to compare the most extreme outlier vs. the average (center) point as shown in Figure 12. Temperature and Dew Point are the main contributors for February 15th (Sample 270). Humidity is relatively more important for the center point (Sample 131).



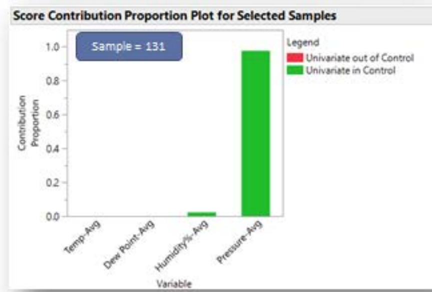


Fig.12 Contribution Analysis Comparison

On February 24th, Governor Abbott called for the Texas grid to be winterized. Many critics believe that the new legislation does not go far enough to prevent another power crisis. Utilities should plan for the future and consider the increased risks associated with climate change instead of looking just at the past.

4. Results and Conclusions

Studied different weather indicators such as temperature, humidity, and dew point. Used Cluster Variables to address the limitations of the general Quantile Outliers tool. Implemented Cluster Variables and the Principle Component Analysis to more efficiently reduce the number of parameters and distinguish the weather pattern. Compared the February 2021 weather situation with the past decade using the Statistical Process Control (SPC) chart, the Heat Map, and the Score plot. Investigated the different contribution proportions of the four most representative variables (temperature, dew point, humidity, and pressure). Temperature and dew point may be the most important weather indicators for future events (instead of humidity and pressure).

Acknowledgements

The authors would like to thank Dr. Charles Chen and Mr. Patrick Giuliano for providing invaluable guidance in the statistical modeling.

References

1. https://www.wunderground.com/history/monthly/us/tx/houston/KHOU/date/2020-2
2. Benford, F. (1938). “The law of anomalous numbers.” Proceedings of the American philosophical society, 551–572.
3. Hartigan, J. A. (1981). “Consistency of Single Linkage for High–Density Clusters.” Journal of the American Statistical Association 76:388–394.
4. Kourti, T., and MacGregor, J. F. (1996). “Multivariate SPC Methods for Process and Product Monitoring.” Journal of Quality Technology 28:409–428.

5. Candes, E. J., Li, X., Ma, Y., and Wright, J. (2009). “Robust Principal Component Analysis?” Journal of the ACM, 58:1–37.
6. Jackson, J. E. (2003). A User’s Guide to Principal Components. Hoboken, NJ: John Wiley & Sons