Algorithm for Identifying Inflection Points in Wind Pressure Data Baseline of Air-Supported Membrane Structures.

Xiangyu Gong, Yanli He*, Xiongyan Li

Faculty of Architecture, Civil and Transportation Engineering, Beijing University of Technology, 100124 Beijing, China

gongxiangyu@emails.bjut.edu.cn (Xiangyu Gong)
yanlihe@bjut.edu.cn(Yanli He)
xylee@bjut.edu.cn (Xiongyan Li)

Abstract. Pneumatically supported membrane structures belong to wind-sensitive structures, and the on-site measured wind pressure data is a very valuable and intuitive data source for pneumatic membrane structures. However, the acquisition of wind pressure data is difficult, costly, and there is very little research on the processing of wind pressure data for pneumatic membrane structures with long duration. In this paper, all wind pressure measurement points on-site of the pneumatic membrane structure were taken, and based on the characteristics of the measured wind pressure data, the data was preprocessed by dividing the data into abrupt change points to remove baseline shifts caused by abrupt changes. The results show that the algorithm proposed in this paper is efficient, and provides flexibility in dividing abrupt change points, laying a foundation for further extracting useful information from the data. This algorithm can also be applied to data processing with similar characteristics in other applications.

Keywords: wind pressure, data processing, baseline shift, air-supported membrane structure.

1 INTRODUCTION

The pneumatic membrane structure belongs to wind-sensitive structures, and research on wind-induced vibration response mainly focuses on three aspects: simulation, wind tunnel tests, and field measurements. Field measurements, as the most direct source of data, directly verify the above two aspects. However, due to the high engineering cost and difficulty in data collection of field measurements, there are relatively few studies on monitoring in this area. Yue Yin[1] et al. first established a health monitoring system for pneumatic membrane structures at Shanghai Jiao Tong University to determine the load and structural response during typhoons. However, the number of measurement points is limited, the measurement time is short, and the data volume is difficult to form a large-scale data selection. Ying, S.[2] regards the wind pressure on large-span spatial

B. Yuan et al. (eds.), *Proceedings of the 2024 8th International Conference on Civil Architecture and Structural Engineering (ICCASE 2024)*, Atlantis Highlights in Engineering 33,
https://doi.org/10.2991/978-94-6463-449-5_36

structures as a stationary non-Gaussian field, and discusses simulation algorithms based on wind tunnel tests.

There are also many studies on baseline drift. Xingxi Shi[3] et al. utilized wavelet transform for preprocessing GPS data. They considered GPS outliers or cycle slips as breakpoints of the signals to be identified, however, these breakpoints still need to be manually identified, which can be a significant workload when dealing with a large amount of measured data. Chen, Z.[4] proposed a deep neural network model based on Empirical Mode Decomposition (EMD-DNN) to address the baseline correction issue by removing the drift trend. Barkauskas D. A.[5] et al. focused on extracting key signals from background noise. Perhaps there is no need to classify the data baseline into discrete and continuous parts, but first identifying the discrete breakpoints of the baseline and then using new technologies such as neural networks for extracting key features will definitely have a more positive impact on the convergence of subsequent neural networks. Chiu[6] et al. solved the numerical algorithm problem by adding a prefixed acceleration pulse and suggested using a third-order polynomial function as the pulse function.

This paper belongs to the wind engineering monitoring project of large-span air-supported membrane structures, with the core objective of obtaining reliable and stable monitoring data for analysis, and ultimately studying the response of large-span air-supported membrane structures through on-site measurements. The approach of data processing in this paper is to divide the baseline shift of the original data into discrete and continuous parts. Firstly, identifying the discrete inflection points to segment the data reasonably, attempting to remove the offset of the discrete part, and then considering further processing of the data. Ultimately, the identification of discrete points and the rational segmentation of the data were achieved. This approach is inspired by the following articles. Lee Eun-Taik[7] et al. proposed a non-baseline damage detection method that is less sensitive to noise. This is similar to the research idea of this article. However, it requires the use of multiple sensors, and manually dividing the known damaged sensor data for comparison with the test data. However, the wind load characteristics in this article are more complex, with most sensors' discrete baseline shifts not occurring at the same time, so it is necessary to grasp the key features of the wind pressure actual measurement data and identify and eliminate the discrete baseline. Wang P.[8] et al. provided a good idea for dealing with the baseline model of a fan engine. The core idea is to find the steady working points in the engine, the more steady states the engine has, the higher the reliability and accuracy of the baseline model. When dealing with baseline shifts in wind pressure data, even if there are problems with the baseline processing of discrete points, as long as the baseline shift points can be accurately found, the time of the wind pressure load can be correctly partitioned as much as possible. Zhang H.[9] et al. proposed a new method to identify the baseline period in a long-term historical dataset. Classifying and verifying data segments based on the discrete points found in this article can further identify usable and reliable data.

2 THE LOCATION OF THE PROJECT.

This article belongs to the wind engineering monitoring project of large-span pneumatic membrane structures. The core objective is to obtain reliable and stable monitoring data for analysis, and ultimately to study the response of large-span pneumatic membrane structures through on-site measurements. The project is located in the third gas-bearing membrane structure coal shed of Yueqing Power Plant in Yueqing City, Zhejiang Province, as shown in Figure 1. The location of the wind pressure measuring point is shown in Figure 2.

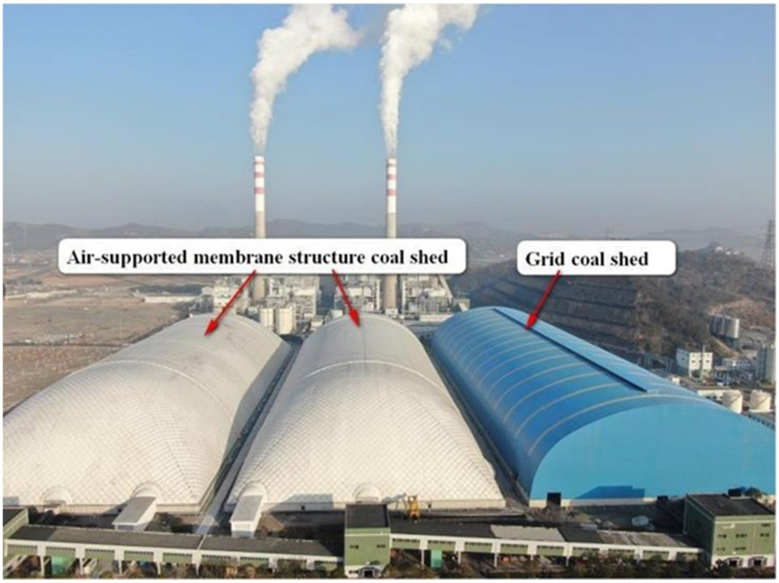

Fig. 1. Relative position of coal shed

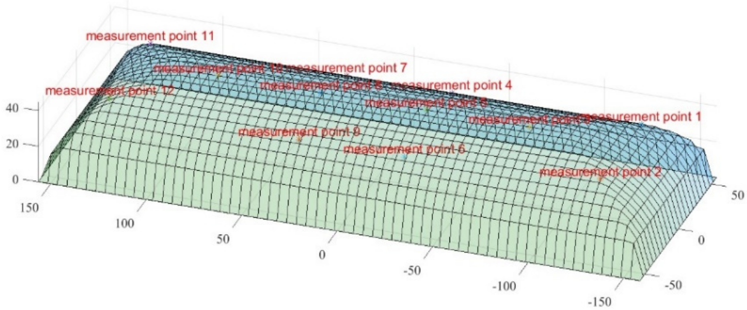

Fig. 2. The location of the wind pressure measuring point

3 THE MAIN PROCESS OF THE ALGORITHM.

First, select the wind pressure measurement data from October to December 2022. The sampling frequency for the first half of these three months is 10 Hz, while for the second half, it is 20 Hz. The purpose of selecting three months of data in this paper is to ensure a sufficient amount of valid data segments can be identified for subsequent research. Wind speed and direction data inherently exhibit strong randomness, and thus, identifying stable and reasonable time segments necessitates a large amount of statistically significant data segments. However, the contradiction arises between the need for a high sampling frequency for selecting research-worthy data segments and the associated challenges in data processing and selection. A higher sampling frequency leads to excessively large data volumes, resulting in significant memory usage and time costs during overall data processing. Conversely, a lower sampling frequency facilitates data selection but may result in the loss of valuable information when studying specific data segments subsequently. Upon observing the overall data, it was found that appropriately downsampling the data does not significantly affect the discrete baseline offset points. Therefore, the overall approach to baseline correction involves initially analyzing the downsampled data, obtaining downscaled baseline offset data comparable to the downsampled data volume, and finally, using interpolation to generate baselines equal in size to the original data, thereby further eliminating baseline offsets from the original data. The specific steps are illustrated in Fig. 3.

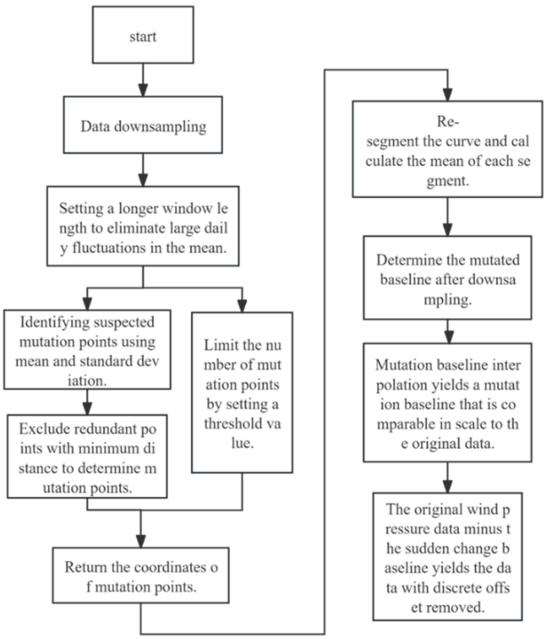

Fig. 3. The process flowchart for Figure 3 baseline removal.

Step one, downsample the data reasonably to facilitate computer processing. Step two, set a relatively long window length and use the moving average method to obtain a smoother data mean curve, eliminating large changes in the mean within a single day. The window should not be too large or too small. An excessively large window can easily lead to information loss, while a too small window may not provide enough smoothing to reflect the baseline characteristics of abrupt change types. Step three, using the same window length, calculate the mean standard deviation and identify suspected mutation points. Step four, set thresholds for minimum distance and quantity. Step five, return the coordinates of the identified mutation points and observe their reasonableness. Step six, if the search for mutation points is deemed reasonable, segment the curve and calculate the mean for each segment. Step seven, starting from the downsampled coordinates, use the mean of each segment as the baseline value to obtain segmented downsampled mutation baselines. Step eight, interpolate the downsampled mutation baselines using the nearest neighbor interpolation method to obtain baselines of the same scale as the original data. Step nine, subtract the baseline from the original wind pressure data to obtain data with removed discrete offsets and the starting points of each data segment.

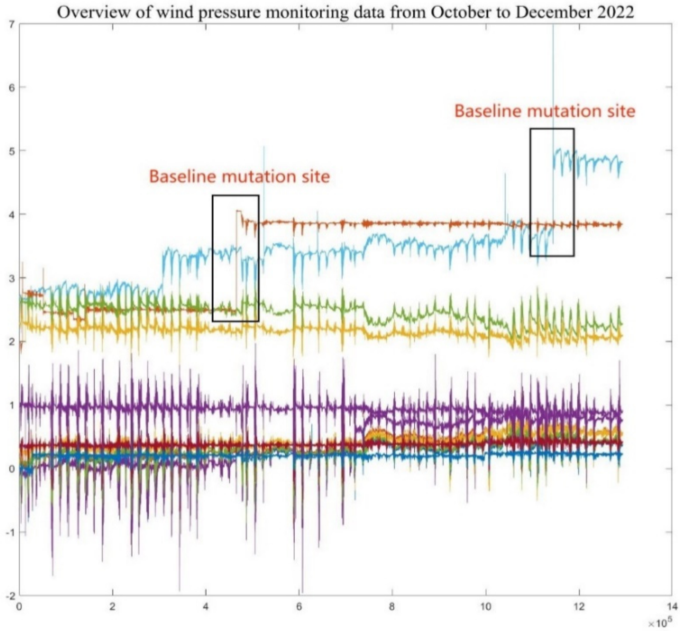

Fig. 4. Original wind pressure data

We have obtained the raw data for 12 measurement points for the months of October to December 2022, as illustrated in Fig. 4. In this study, we gradually tested windows ranging from 1% to 5% of the total data length and observed the results. Ultimately, we selected a window length of 5% to obtain the smoothed data as shown in Fig. 5. In the three months of actual wind pressure data, this window size is slightly less than a day in duration. One assumption we made is that the interval between mutation points is less than a day, as evidenced by the case of Point 9, which had baseline mutations occurring at intervals exceeding a month. Additionally, we found that the actual wind pressure data exhibits a noticeable periodicity, roughly corresponding to a one-day interval cycle. The reason for this periodicity has not yet been analyzed. Choosing a window of this size can to some extent mitigate the interference of large changes in daily data averages on mutation points, thereby enhancing the robustness of the algorithm.

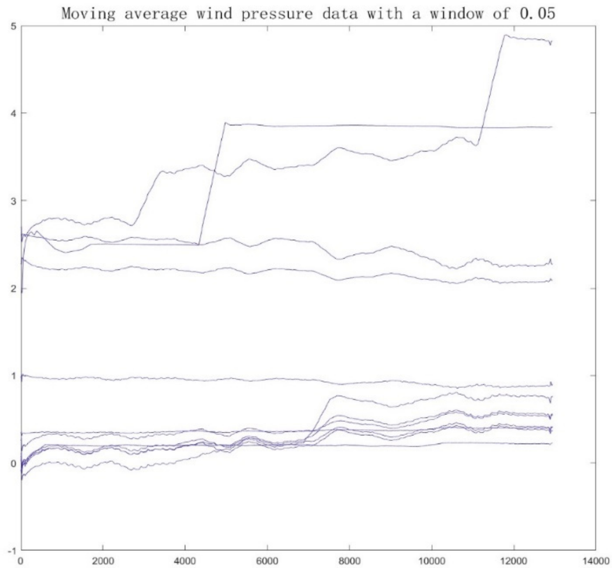

Fig. 5. Wind pressure mean data after sliding average

4 RESULTS AND DISCUSSION OF THE ALGORITHM

This article classifies data segments from few to many through two constraints, one being the limitation on the number of breakpoints and the other being the limitation on the threshold. Among them, the limitation on the threshold plays a basic discriminative role, if the data in the same segment is significantly higher than the standard deviation, it is considered that there is a mutation-type offset here. The limitation on the number of breakpoints is because the range of wind pressure fluctuations at 12 measuring points is different. Only using the threshold constraint may result in some breakpoints between measuring points with small fluctuations difficult to find, while lowering the threshold will make measuring points with larger fluctuations show breakpoints everywhere. The limitation on the number of breakpoints and the setting of the threshold can be gradually adjusted to an appropriate value. As shown in Fig. 6, if the data itself has minimal fluctuations, it can be assumed that there are no significant pressure jump points and can be directly used for subsequent research. As shown in Fig. 6 to 8, the final results obtained in this study are relatively conservative after removing the discrete baselines, with a maximum number set to 3 and a threshold set to 0.084. There are a total of 12 measurement points, among which points 1, 2, 3, 5, 7, 8, 10, 11, and 12 are considered to have no significant baseline shift due to their very small variations. Points 4, 6, and 9 show slight shifts, with point 9 being the most unusual among all measurement points, as visible abrupt changes have been identified. For the sake of brevity, this paper only presents the baseline identification and elimination figures for three measurement points.

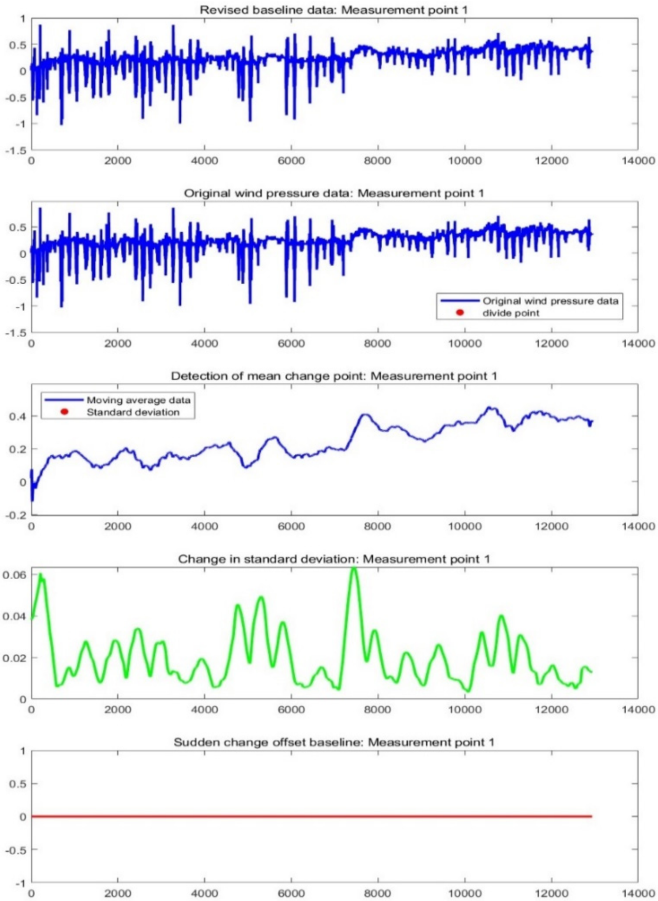

Fig. 6. Standard deviation detects mean shift at test point 1.

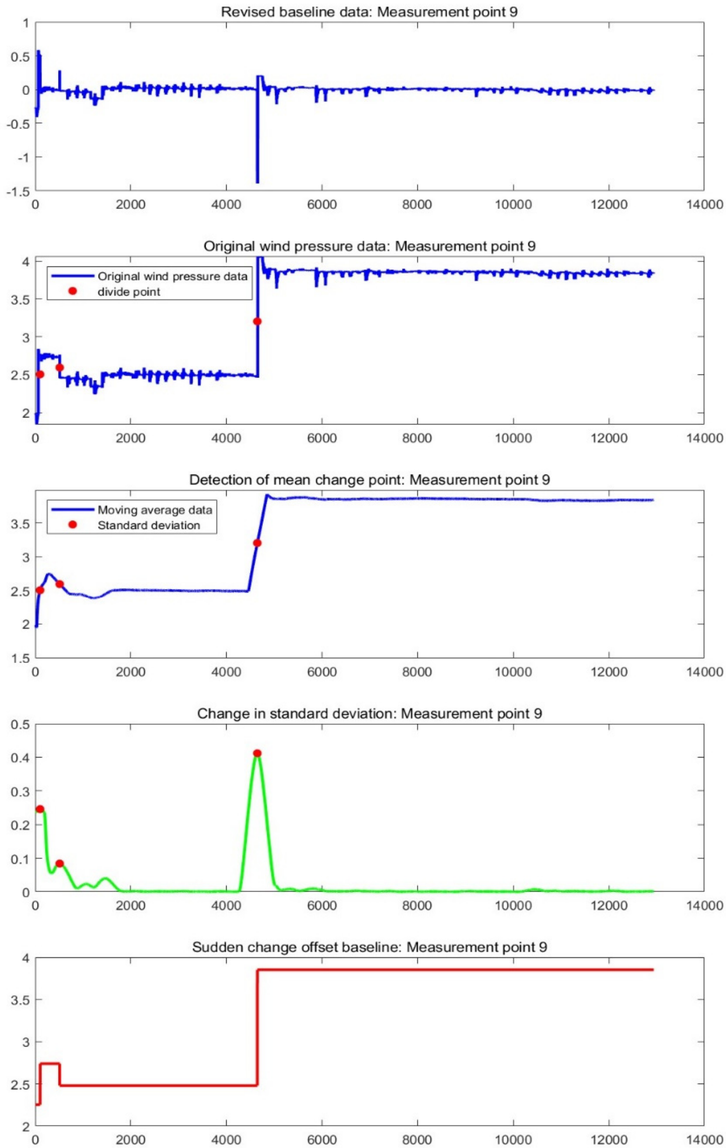

Fig. 7. Standard deviation detects mean shift at test point 9.

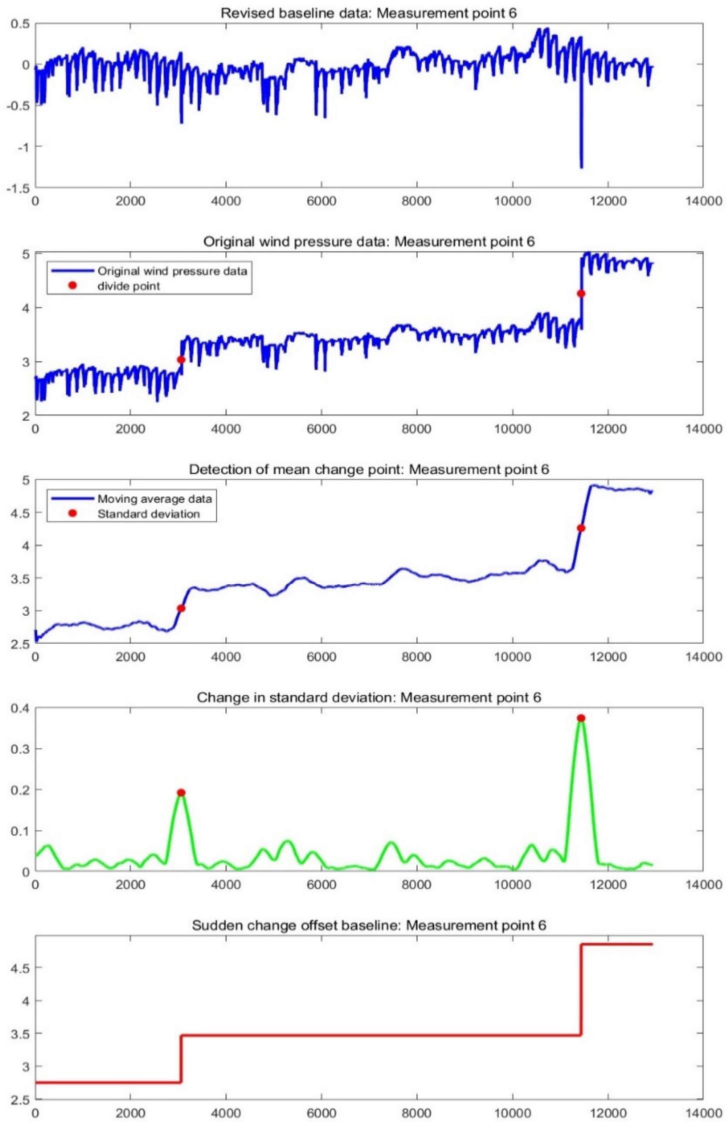

Fig. 8. Standard deviation detects mean shift at test point 6.

If a more aggressive strategy is adopted, setting the threshold at 0.05 and limiting the number to only 8, the results shown in the figure will be obtained, as shown in Fig.

9. We can see that, except for measurement points 3, 7, 10, and 11, there is a baseline shift in most of the points. Points 1, 2, 4, and 5 have the same abrupt change in baseline position, all contained within the position of point 6. We can consider that the cause of the mutation at this time should be the same.

Therefore, setting a smaller threshold and a larger quantity for the more aggressive removal of baseline mutations makes it easier to identify baseline mutation time points that cause overall structural changes. Using such data for structural classification is more reasonable and also makes it easier to find and filter data in the steady state, reducing the interference of data changes in discrete features on subsequent research. On the other hand, setting a larger threshold is more conservative, and the data points of the discrete parts found are more accurate, making it less likely to mistakenly identify discrete data points. Because subsequent research relies on the continuity of data, mistakenly identifying discrete points and then smoothing the data will introduce new discreteness and lead to misjudgments in the results.

After reasonable data segmentation, the causes of abrupt offset formation in the overall data can be analyzed by combining acceleration, displacement, and internal pressure data in the monitoring of pneumatic membrane structures. Subsequently, the threshold can be set incrementally from large to small, and the number of abrupt points can gradually change from few to many, combined with the data of subsequent structural responses to further assist in determining the rationality of identifying abrupt points in this paper. Each stable period found can be explored for its characteristics using a similar continuous data processing approach.

5 CONCLUSIONS AND PROSPECTS

This paper presents a relatively free algorithm for detecting discrete baseline shifts in wind pressure load data. This approach successfully eliminates discrete baseline shifts in wind pressure data for large-span air-supported membrane structures. Taking three months of data as an example, based on the maximum number of discrete points and standard deviation threshold set in this paper, the effectiveness of the proposed algorithm was verified, providing a basis for further research on measured wind engineering data. The main conclusions of this paper are as follows:

1. The selection of window size: a window length slightly larger than half a day but less than a day can effectively eliminate the interference of wind pressure data value mutations with a daily cycle.

2. By using two indicators, threshold and maximum number, baseline shift points can be identified effectively, providing a high degree of freedom in selection.

3. A conservative selection of quantity and threshold can better identify shift points, while a more aggressive selection can result in more stable data segments.

After segmenting the data using mutation points, it is possible to determine the consistency of different mutation points over time to ascertain whether there is a common cause for the current mutation. Subsequently, the segmented data can also be further processed to remove continuous baseline offsets, or applied to machine learning to capture key features. As machine learning itself requires data stability, normalizing the data after eliminating mutations will have a positive effect.

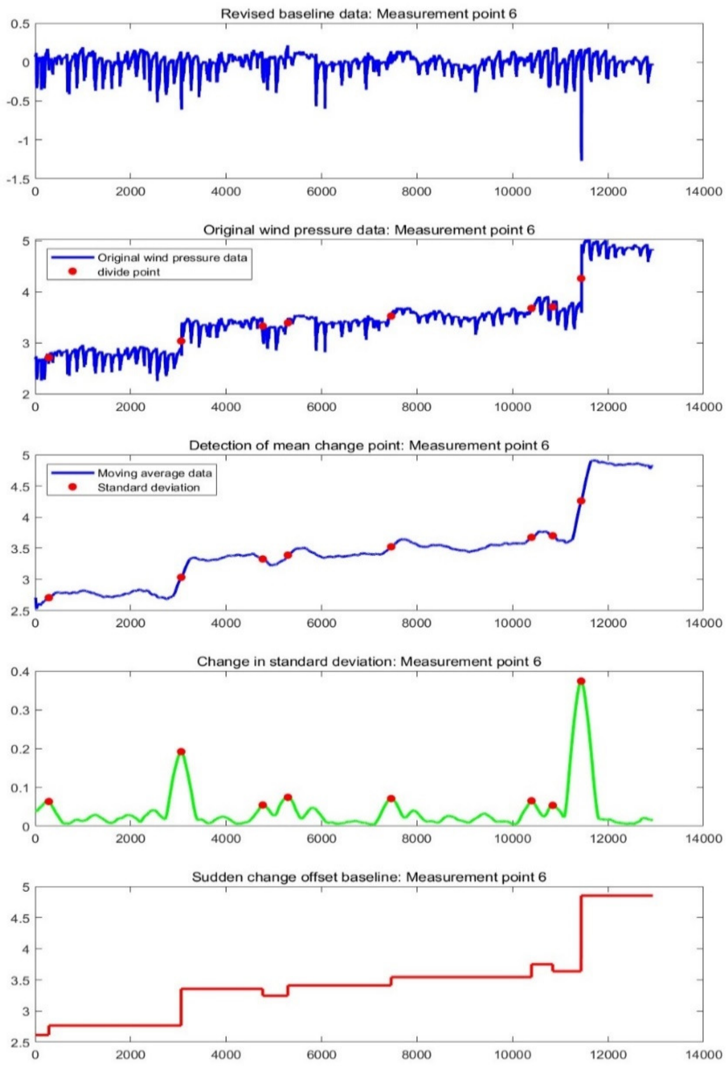

Fig. 9. Standard deviation detects mean shift at test point 6, limit 8 points

REFERENCES

1. Yue Yin, Wujun Chen, Jianhui Hu, Bing Zhao, Qin Wang, In-situ measurement of structural performance of large-span air-supported dome under wind loads, Thin-Walled Structures, Volume 169, 2021
2. Ying, S, Ning, S, Yue, W, Multivariate stationary non-Gaussian process simulation for wind pressure fields. EARTHQUAKE ENGINEERING AND ENGINEERING VIBRATION. Volume15, Issue4, Page729-742, 2017
3. Xingxi Shi, Chunxia Zhao, Longhe Zhu, Weimin Ju, Shuhe Zhao, Remotely Sensed Data and Information - Data pretreatment in GPS baseline solution based on wavelet, SPIE Proceedings [SPIE Geoinformatics 2007 - Nanjing, China (Friday 25 May 2007)] Geoinformatics 2007 doi:10.1117/12.761293
4. Chen, Z., Fu, J., Peng, Y., Chen, T., Zhang, L., & Yuan, C. (2021). Baseline correction of acceleration data based on a hybrid EMD–DNN method. Sensors, 21(18), 6283. doi:https://doi.org/10.3390/s21186283
5. Barkauskas, D. A., Rocke, D. M. (2010). A general-purpose baseline estimation algorithm for spectroscopic data. Analytica Chimica Acta, 657(2), 191–197. doi:10.1016/j.aca.2009.10.043
6. Chiu, H.-C. (2012). A Compatible Baseline Correction Algorithm for Strong-Motion Data. Terrestrial, Atmospheric and Oceanic Sciences, 23(2), 171. doi:10.3319/tao.2011.10.25.01(t)
7. Lee Eun-Taik, Eun Hee-Chang. Damage identification through the comparison with pseudo-baseline data at damaged state. Engineering with Computers, 32(2), 247–254, 2016 doi:10.1007/s00366-015-0413-2
8. Wang P., Liu K., Tang Z. Turbofan Engine Baseline Model Extraction Based on FDR Data. 2019 Chinese Control And Decision Conference (CCDC), 2019 doi:10.1109/ccdc.2019.8833387
9. Zhang H., Albin S. L., Wagner S. R., Nolet D. A., Gupta S. (2010). Determining Statistical Process Control Baseline Periods in Long Historical Data Streams. Journal of Quality Technology, 42(1), 21–35. doi:10.1080/00224065.2010.11917804