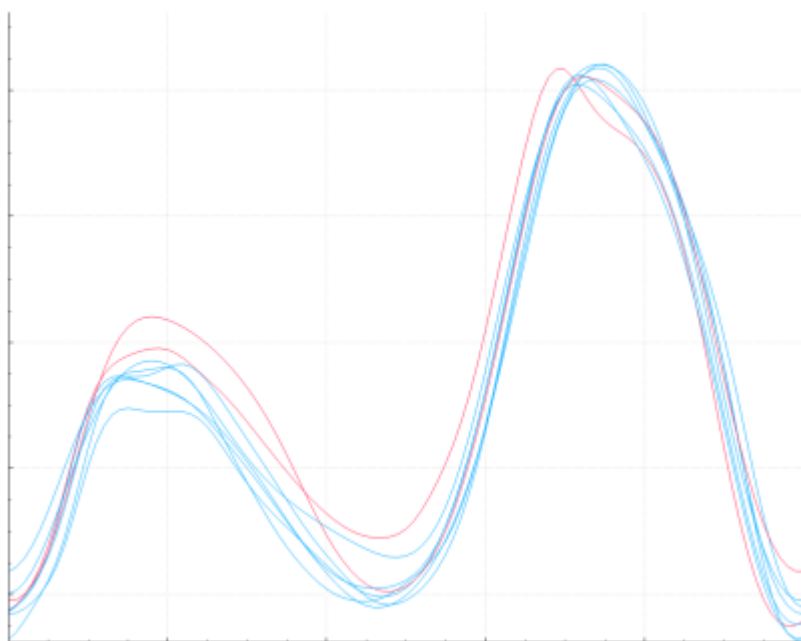


Anomaly Detection

Sift provides a systematic approach for identifying and handling outliers in biomechanical data. The general workflow is:

1. **Data cleaning:** exclude traces that clearly do not represent real movement, before any statistical analysis.
2. **Outlier detection:** apply one or more detection methods to identify less obvious anomalies.
3. **Act on results:** exclude detected outliers, or treat them as traces of interest for further investigation.



Anomaly Detection Using Dynamic Time Warping

Data Cleaning

Before running any outlier detection, it is worth removing obvious data quality issues (e.g. missed markers, incorrect force plate assignments, or trials where the participant did not perform the task as expected). These drastic cases can skew the representation of “normal” and affect the reliability of downstream detection methods.

Follow the [Clean your Data](#) tutorial, also available as [Sift Tutorial Video 2: Cleaning Data](#).

Outlier Detection Methods

Mahalanobis Distance Test and SPE

Two complementary methods for detecting outliers in PCA space.

- **Mahalanobis Distance** measures how far a data point is from the centre of the distribution, accounting for the covariance between dimensions. Use this to find points that appear close to the mean data but do not follow its general trend.
- **Squared Prediction Error (SPE)** measures how well a data point is reconstructed by the PCA model. Use this to find points that are poorly represented by the retained components. These points contain variance that is not captured by the PCA model, even if they appear close to the group mean in PC space.

When to use: Use these methods to detect global outliers in the workspace scores after running PCA.

- Learn more: [Mahalanobis Distance Test](#), [SPE](#)
- Tutorial: [Mahalanobis Distance and SPE Tests](#)

Local Outlier Factor

Local Outlier Factor (LOF) detects outliers by comparing the local density of each point to that of its neighbours. Points in sparse regions relative to their surroundings are flagged as outliers. In Sift, LOF has been attached the PCA module.

When to use: Use this method to find local outliers in the PCA workspace scores, and the results contain clusters of varying density.

- Learn more: [Local Outlier Factor](#), [Local Outlier Factor Dialog](#)
- Tutorial: [Local Outlier Factor Tutorial](#)

K-Means

K-means groups PC scores into clusters based on Euclidean distance. Points that lie far from their cluster centroid are candidates for outliers.

When to use: Use this method when the data is expected to naturally fall into distinct groups in PC space, and to identify points that do not fit well within any group. Requires specifying the number of clusters in advance.

- Learn more: [K-Means Dialog](#)
- Tutorial: [Run K-Means](#)

Other Methods of Outlier Detection

Dynamic Time Warping

Dynamic Time Warping (DTW) measures the similarity between two time series traces by finding the point-to-point alignment that minimises total distance. The result is a cost function (the higher the cost, the less similar the two traces are). DTW does not require PCA.

For anomaly detection, each trace is compared to every other trace in the dataset. Traces whose median distance from all others exceeds a specified threshold are flagged as anomalies.

When to use: Use this method to detect outliers based on signal shape, without first running PCA.

- Learn more: [Dynamic Time Warping](#)
- Dialog: [Dynamic Time Warping Dialog](#)

Excluding Detected Outliers

Once outliers have been identified, they can be flagged for exclusion without deleting the underlying data. See [Exclude Data in Sift](#) for a full overview.

Method Comparison

Method	What it detects	When to use it	Requires PCA?
Mahalanobis Distance and SPE	Points that are global outliers in PC space	Detecting global outliers in PC scores after running PCA	Yes
Local Outlier Factor	Points that are unusually isolated from their neighbours in PC space	Outliers are suspected within subgroups, not just across the whole dataset	Yes
K-Means	Points that do not fit well within any natural grouping	A hypothesis exists about how many subgroups are in the data	Yes
Dynamic Time Warping	Traces dissimilar in shape to others in the dataset	Shape-based outlier detection without PCA	No

From:

<https://www.wiki.has-motion.com/> - **Wiki Documentation Site**

Permanent link:

https://www.wiki.has-motion.com/doku.php?id=sift:sift_overview:anomaly_detection

Last update: **2026/03/31 15:35**

