

Clean_your_Data

This tutorial will show you how to use Inspect3D as a data cleaning, or quality assurance, tool. You will learn how to check for and correct faulty force assignments from all files at once, thereby removing any artifacts from the data. This is particularly useful for lab managers (or supervisors) who may not be familiar with the raw data, or collaborators that were not involved in the collection process.

Data

This tutorial uses overground walking data from four subjects. The subjects walked at three different speeds; slow, normal and fast. The data was analyzed using a pipeline which included an automatic gait event detection command. [Link model based items](#) (including right and left ground reaction forces) were created using a pipeline. This data is reflective of a typical outcome from batch data analysis in [Visual3D](#).

These data files can be downloaded using this link: [Four Subjects Walking Data Set](#).

Loading the library



1. Click  **Load Library** in the [toolbar](#) to open the Load Library dialog.
2. Click **Browse** and select the folder where the CMZ files are stored.
3. Click **Load** button to import the data.

This step selects the path to the data you are using. If you intend to modify the data, e.g., by correcting invalid assignments, then your data must be saved in a directory where you have read and write access. This tutorial uses a folder located on the desktop. The data folder must contain the .cmx file associated with every [.cmz](#) file, since without both file types Inspect3D will be unable to load the files. Each .cmz file has multiple .c3d files with the different walking speeds identified using Tags.

NOTE: After the library path is set, files from the selected folder and sub-folders will be loaded and parsed. Inspect3D will show a summary of the loaded files and what they contain in the Load Library dialog. Examine this summary to make sure that the correct files have loaded.

Defining queries and calculating groups

1. Click on the  **Query Definitions** icon on the [toolbar](#) and then click on  **Query Definitions** to open the Query Builder dialog.
2. To create a new query definition, click the  green addition button in the empty **Queries** list.
 - 2.1. Type GRF in the **Query Name** text box in top-right and click **Save**.



3. While the GRF query is selected, click the  green addition button in the empty **Conditions** list.

3.1. Type R_GRF in the **Condition Name** text box in the top-right.

3.2. There are now three tabs that need to be completed in order to define the sub-group.

3.2.1. **Signals:** This tab allows the user to select specific signals according to their names in the [Visual3D Data Tree](#). For this tutorial, select the signal: TYPE - LINK_MODEL_BASED, FOLDER - ORIGINAL, NAME - R_GRF, and COMPONENT - Z.



3.2.2. **Events:** This tab allows the user to specify the desired event sequence to extract data from. For instance, the right gait cycle could be extracted using the event sequences RON, ROFF or RHS, RHS. For this tutorial select the RON and ROFF events and leave the default normalization values of 101 for the number of points and cubic for the spline type.



3.2.3. **Refinements:** This tab allows the user to refine their selected signal using two separate and distinct methods: according to the signals contained within the [CMO files](#), or according to Tags. For this tutorial, Refine using tags by leaving the **Use AND Logic** checkbox unchecked and selecting the SLOW, NORMAL, and FAST tags.



3.3. Click **Save** to create the condition.

4. For this tutorial, define a second condition within the same group to account for the left side. This can be done by modifying the existing sub-group:

4.1. While the GRF group is selected, click the  green addition button in the **Sub-Groups** list and type L_GRF in the **Sub-Group Name** text box.

4.1.1. In the Signals tab, change the NAME to L_GRF and leave the other parameters as they are.

4.1.2. In the Events tab, select the event sequence LON and LOFF.

4.1.3. Leave the Refinements tab as it was.

4.2. Click **Save** to create this second condition.



You should now see one query (GRF) in the **Queries** list and two conditions (R_GRF and L_GRF) in the **Conditions** list. At this point in the tutorial these definitions have been created but they have not been applied to the signals in the loaded library. To do this, click on **Calculate All Queries** (or **Calculate Selected Queries**, since there is only one group in this tutorial).

Once the group is calculated you can close the Query Builder dialog, it is now time to examine the results visually by plotting them.

Visualizing and exploring your data

1. With the groups calculated, the specific traces that have been extracted from the loaded library's .c3d files can be explored. The group name defined in the previous step is shown in the **Groups** section (see 1). When the groups are clicked, the workspaces associated with the selected group(s) are shown (see 2). The icon beside a workspace indicates its colour in the **Queried Data** plots. A check mark indicates that the workspace has had none of its traces excluded; an X indicates that at least one trace has been excluded.



2. Select all of the workspaces either by using the **Select All Workspaces** check box or by selecting them all in the **Workspaces** list. Then ensure that the **Plot All Sequences** check box is checked and click **Refresh Plot**.

3. When Inspect3D tries to plot **missing data**, your y-axis will go to -999 999.0 and the axis scale will cause the data to be compressed into lines. To avoid this, open the  **Options** dialog and deselect the **Show data not found** checkbox in the **Plotting options** section.

3.1. Update your plot by clicking **Refresh Plot**; it should now resemble the graph below.



4. Use your cursor to select only lines on the graph that you wish to inspect. Click on single traces in order to examine individual curves without the other curves 'polluting' the graph. When a trace is clicked, a tooltip will appear to describe the source of the trace in terms of both Inspect3D's groups and workspaces and the underlying .c3d files.



Verifying raw data and excluding traces

Once the queried traces have been plotted, it is now possible to begin cleaning the data. In this example, in particular, it is evident that some incorrect forces assignments exist.

1. In order to exclude these incorrect assignments, select any lines that may be a result of a 'bad' event. For the obvious 'bad' events, it is possible to click and drag over multiple traces on the plot to select them simultaneously.

2. Right-click and navigate to Exclude, at the very bottom, and click Exclude Trace (raw data).

3. Verify that your desired traces have been properly excluded.

3.1 When traces are excluded, two notable differences should appear in the Queried Data subwindow. First, the previously selected, but now excluded data should not be visible on the graph. Second, the **Workspaces** widget will indicate when traces have been excluded. Specifically, an X should replace the original check mark beside each workspace where signal traces have been excluded, two horizontal bars should replace the check mark beside each .c3d file with excluded traces, and a red X should replace the checkmark beside each excluded trace. Additionally, the fraction of traces excluded for each workspace is displayed on the right-hand side of the **Workspaces** widget: in this case all of the workspaces contained poor quality events. If unintentional exclusion occurred for a

subject, you can right click on the subject and select 'Re-Include Data'.



4. If you want to visualize the data you have excluded alongside the remaining data, go to the [Show Options](#) dialog and select the “Show excluded data” option. You can choose to use a specific line style for excluded data and specify this style in the [Line Styles](#) tab.



5. 

Traces can also be examined more closely before deciding to include or exclude them. 5.1 Select the traces of interest by clicking and dragging on the plot.

5.2 Click the  **Show Animation** button.

5.3 The Show Animation dialog can display animations for those selected traces with enough data in the .cmz file to permit a recreation. Each of the selected traces can be selected from the drop-down box near the top of the dialog. This view can be used to determine what (if any) issue exists in the data and whether this merits exclusions. Excluded traces are removed from the drop-down box within the dialog as well as the Queried Data subwindow.

6. Once all of the desired exclusions have been made, the original data is ready to be updated.

Updating the original data

1. 

Now that the incorrect force assignments have been identified in our data set, it is possible to update the original [CMZ](#) data by clicking the  **UpdateCMZs** item in the application's tool bar. This opens a dialog that provides different options for processing excluded data. 2. The **Excluded Traces** options allow the user to add a BAD event to the excluded traces. This option is helpful if the intention is to [tag](#) the bad data, re-open the original data, and make manual corrections one trace at a time.

3. Alternatively, the **Update Force Assignments** option actually removes the force assignment from the original data. This option should be used if the user does not intend to revisit the original data after the modifications have been made (note: this second option changes the underlying data. Re-performing this tutorial with the newly saved data will produce different results).

4. Select both **Add Event to Exclude Signals** with the default event name “BAD” and **Update Force Assignments using excluded Query data from:** with “GRF”. Then click **Update CMOs**.

5. This results in Inspect3D opening each .cmz file in the background, adding a BAD event for the excluded traces, and removing the force assignments. This process can take some time depending on the .cmz size, but you can further explore your data (or get a coffee) as it is occurs. When the update is complete, the check mark icons will return beside each workspace in the **Workspaces** widget and the original data will be modified.

Recap

In this tutorial you learned how to use Inspect3D as a data cleaning tool for quality assurance purposes. Here, the process of loading data, defining queries, excluding data, and updating the original data are described.

Navigate Tutorials

Previous: [Tutorial 1: Load and View Data](#)

Inspect3D Getting Started Tutorials

Next: [Tutorial 3: Perform Principal Component Analysis](#)

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

Permanent link:
https://has-motion.com/wiki/doku.php?id=inspect3d:tutorials:clean_your_data&rev=1720027572

Last update: **2024/07/03 17:26**

