Generic Calculations of Curve Slopes/Endpoints

This platform can be utilized for calculation of several metrics from response data over some measure of time. Originally intended for colorimetric development, the platform now works for any kind of numerical response over a unit of time (click here for the list of metrics and how they are calculated).

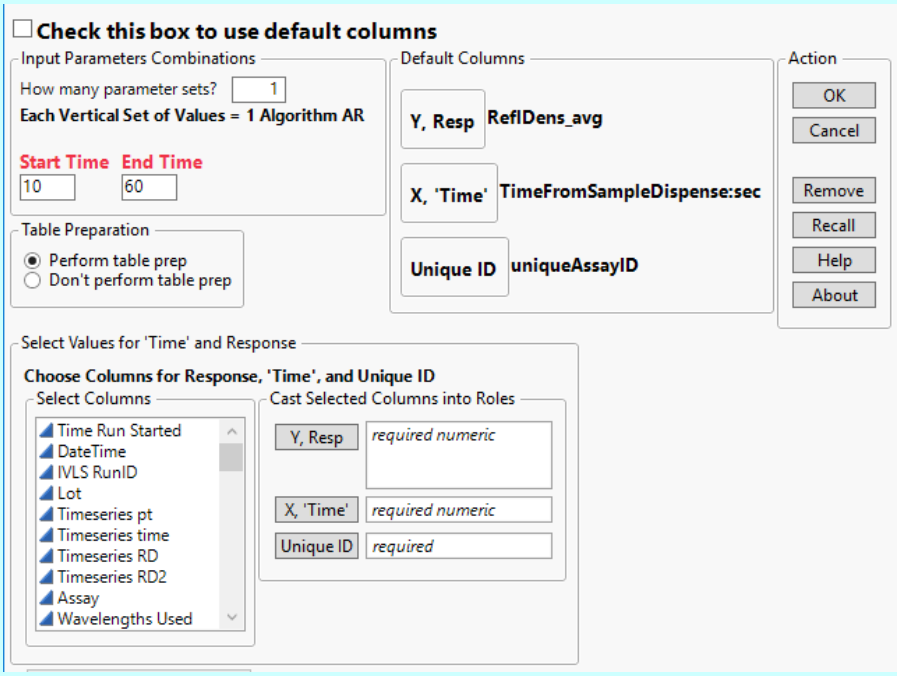

Figure 1.

Within the dialog (Fig 1), the user will have several options. These options help the user decide how they want to construct their metrics as well as the columns to use. The top of the dialog has a checkbox with large, bold, text to use default column names (listed in the Default Columns panel box). If this is checked, there will be no area to select columns to use (Fig 2).

Below the check box are the input parameters for calculation of metrics. This platform allows for the use to input as many metrics as they want for a mixture of all the metrics they choose (Fig 3). The metrics are chosen based on how the user inputs information into the Start Time and End Time (a numeric End Time is required for all metrics).

Slope – A numeric Start Time and End Time

Note: A failed Start Time value (non-'a' or 'r' text, for example) will result in a Slope metric with Start Time = 0

Endpoint – Any text in Start Time (no numbers) with a numeric End Time

Average Endpoint – The letter **"a"** preceding the Start Time number along with a numeric End Time

Range - The letter **"r"** preceding the Start Time number along with a numeric End Time

Note1: Any Start Time that doesn't have a point prior to it (which is in positive 'Time') will be extrapolated to from the 1st two points.

Note2: Any End Time that doesn't have at least one actual point after that 'Time' will not be extrapolated and will result in a value of -999

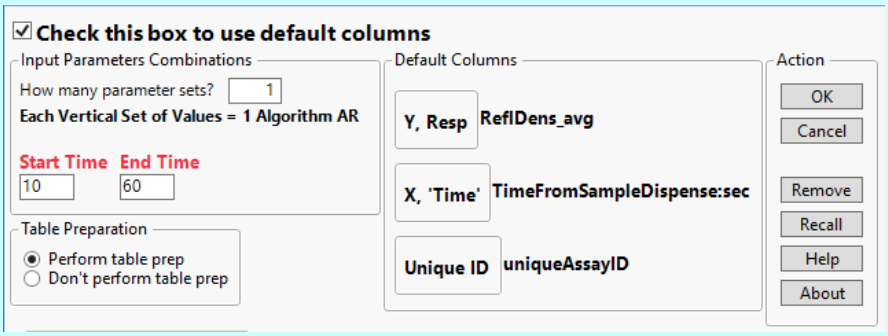

Figure 2.

Beneath the Input Parameters is the option for the user to apply 'table preparation' to the data. Table preparation is an included function which will take the users data and convert column names to commonly used column names across sources so data can be easily understood and combined across data sources, experiments, etc. Table Preparation is highly recommended for any data created from the Catalyst Dx™ or Catalyst One™ platforms and is automatic when using other tools. Table preparation will also parse out flagging notes from log outputs if available.

Note: Table Preparation will also organize the columns for ease of access to the most used ones, so if the table preparation has already been done, selecting the 2nd radio button will guarantee any column organization put in place by the user.

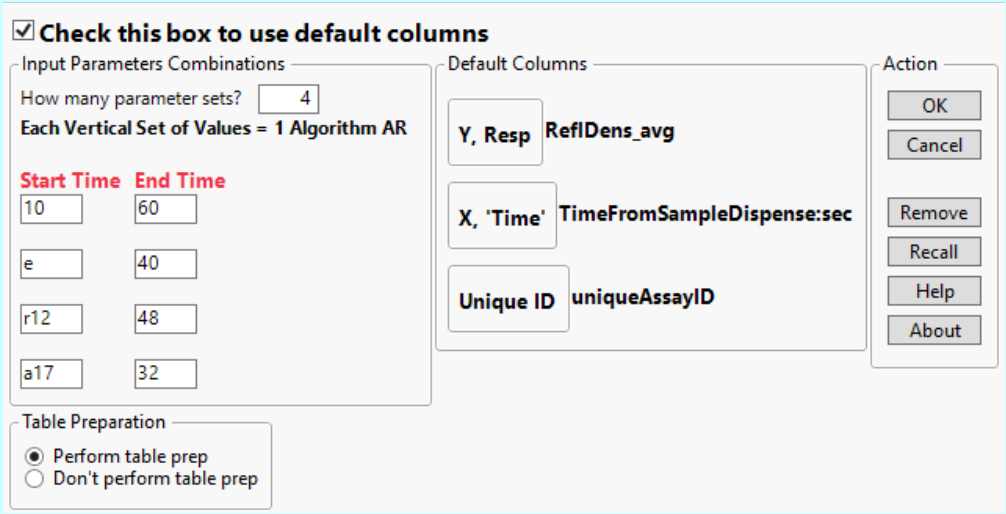

Figure 3.

The Select Values for the 'Time' and Response panel box gives the user the ability to select their own columns to use for analysis should the defaults not be present or applicable to the data being described. If the user choses to pick their own columns, all three need to be filled in – with one column only in each section.

Y, Resp – The response for the data

X, 'Time' – The surrogate for time. Time is often used, but non-time representations of time can also work (e.g. frame count, read point, etc.) if they are numeric/continuous
NOTE: If a given run has a first-row-time = 0, the algorithm will return -999 for that run. The first row in the run must be non-zero. Accounts for Catalyst errors where runs contain 1+ rows starting with time = 0 seconds.

Unique ID – A fully unique identifier for each time/response data in the data set. It can be created by the user or created by Table Preparation
Note: Depending on what is available in the data, Table Preparation won't be able to make a completely unique ID

The panel box on the far right has all the action options for the dialog.

OK – Takes the user inputs and analyzes the data with the resulting outputs (described below)

Cancel – Closes the dialog
Remove – Removed the selected columns from the column boxes
Recall – Fills in all the areas with the last successful execution of data so the user doesn't need to reenter man different parameters (for example)
Help – Opens this document
About – Lists the version of the platform along with the author

Outputs:
The tool will calculate and output your data in one of two methods: a new table with curves / calculations; additional columns in the current table with calculations.

If the data is considered untouched or 'raw' the tool will copy it and do all work on this copy. This allows the user to maintain full control on raw data. The script determines if the data is 'raw' with a tag contained in the filename and is a common output for other tools (including this one). Any file containing "_progCurves" in the filename will be treated as not 'raw' and calculations will be done within this table. If this tag is not present, a new table will be created which contains this tag so the user can save their 'raw' data, still untouched.

Calculated columns are added to the far right in the table and labeled 'Slope', 'Endpoint', 'Range' and 'AvgEP' based on the metric along with the time range covered.

Whether a new table is created or the existing one is used, a table script "Make justResults" which turns the "progCurves" data into a table containing just the results for the runs (one row per run). This script will also append "_justResults" onto the table name to encourage the separation between these two types of files.

Calculations:
Calculations are done based on interpolation (or limited extrapolation) to the exact times listed in the Start and End Times.

Slope – The regression slope for all the points between the Start and End time inclusive of the two points interpolated to (see example below)
Endpoint – The single interpolated value to the exact End Time
Average Endpoint – The average for all the points between the Start and End time inclusive of the two points interpolated
Range – The range of all the points between the Start and End time inclusive of the two points interpolated

<u>Slope Example:</u>
If there are reads at 1, 2, 3, 4, and 5 seconds and you want an AR based on values from 1.3 seconds to 4.2 seconds the macro will use the 1 and 2 second reads to linearly interpolate to a 1.3 second value along with the 4 and 5 second reads to interpolate to a 4.2 second read. Then, the macro will use the 1.3, 2, 3, 4, and 4.2 second read and perform a slope regression to obtain the final AR value.