Command Reference: ReadDelimitedFile()
Read time series from a delimited file
Version 09.06.03, 2010-04-05
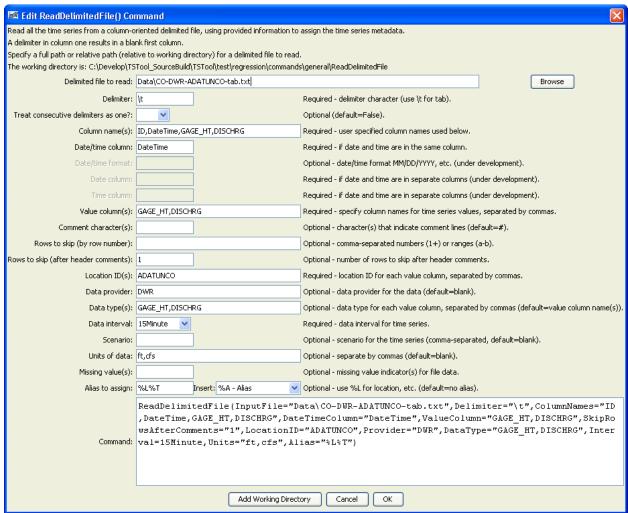
The `ReadDelimitedFile()` command reads one or more time series from a column-oriented delimited file. This command is useful for mining data from the web (see also the `WebGet()` and `FTPGet()` commands) and processing spreadsheets that have been saved as comma-separated-value (CSV) files. The command processes three main types of information:

1. Comments in the header (before data) and embedded in data records (e.g., because bad data values were commented out).
2. Data records, in column format, containing date/time strings, data values, and other information.
3. Metadata, such as station identifiers, data types, units, and interval.

Although the file being read may include metadata within column headings and data records, the information can be difficult to extract because of formatting. For example, column headings may include the data type as "Precipitation\n(in)" (where \n indicates a newline). Consequently, the command requires metadata to be assigned via command parameters in order to ensure robust data handling. Rather than try to automatically determine the data type and units from the column heading, the values can be assigned with the `DataType` and `Units` parameters. Additional functionality may be added in the future automate metadata discovery. An example of a delimited file is shown below (this file can be automatically retrieved using the `WebGet()` command – see the example command file below):

```
#---------------------------- Provisional Data ------------------------------
#This system is maintained by the Colorado Division of Water Resources.
#Contact: Colorado Division of Water Resources (303) 866-3581
#
#All data presented on the Colorado Surface Water Conditions web site are
#provisional and subject to revision. Data users are cautioned to consider
#carefully the provisional nature of the information before using it for
#decisions that concern personal or public safety or the conduct of business
#that involves substantial monetary or operational consequences.
#
#Data is returned in TAB delimited format. Data miners may find help on automating
#queries and formatting parameters at http://www.dwr.state.co.us/help
#
#Gaging Station: ALVA B. ADAMS TUNNEL AT EAST PORTAL NEAR ESTES PARK (ADATUNCO)
#Retrieved: 3/30/2010 03:04
#----------------------------------------------------------------------------
Station Date/Time        GAGE_HT (ft)    DISCHRG (cfs)
ADATUNCO        2006-10-01 00:00        2.34    225
ADATUNCO        2006-10-01 00:15        2.34    225
ADATUNCO        2006-10-01 00:30        2.35    227
ADATUNCO        2006-10-01 00:45        2.34    225
ADATUNCO        2006-10-01 01:00        2.34    225
… etc …
```

The following dialog is used to edit the command and illustrates the syntax for the command.

ReadDelimitedFile() Command Editor

The command syntax is as follows:

ReadDelimitedFile(Parameter=Value,…)

Command Parameters

Parameter	Description	Default
InputFile	The name of the delimited input file to read, surrounded by double quotes to protect whitespace and special characters. Global property values can be used with the syntax `${PropertyName}` (see also the `SetProperty()` command).	None – must be specified.
Delimiter	The delimiter character(s) that separate columns.	None – must be specified.

`TreatConsecutive DelimitersAsOne`	Indicate whether consecutive delimiter characters should be treated as a single delimiter, for example, when multiple spaces are used to line up columns.	`False` (columns are separated by a single delimiter character)
`ColumnNames`	The user-specified names for columns in the file, used to ensure that column headings in files are properly interpreted. These names are used in other parameters to specify columns in the file. Separate column names with commas.	None – must be specified.
`DateTimeColumn`	The column matching a value in `ColumnNames`, which indicates the date/time column in the file.	None – must be specified.
`DateTimeFormat`	The format for date/time strings in the date/time column.	This parameter is current under development – the format is automatically determined in most cases.
`DateColumn`	The column matching a string in `ColumnNames`, which indicates the date column in the file.	This parameter is currently under development.
`TimeColumn`	The column matching a string in `ColumnNames`, which indicates the time column in the file.	This parameter is currently under development.
`ValueColumn`	The column(s) matching a string in `ColumnNames`, which indicate the data value columns. Separate column names with commas.	None – must be specified.
`Comment`	Character(s) that if found at the start of lines in the file, indicate that the line is a comment. The characters are interpreted individually (e.g., `#$` indicates that lines starting with `#` or `$` will be treated as comments).	`#`
`SkipRows`	Indicate absolute rows (1+) in the file to skip, using single numbers and ranges a-b, separated by commas. Rows are skipped prior to other processing.	No rows will be skipped.
`SkipRowsAfter Comments`	Indicate the number of rows to skip after header comments. Use this parameter to skip column headers prior to the data lines.	No rows will be skipped.
`LocationID`	The location identifier(s) to assign to time series for each of the value columns (or specify one value to apply to all columns).	None – must be specified.
`Provider`	The data provider identifier to assign to time series for each of the value columns (or specify one value to apply to all columns).	No provider will be assigned.
`DataType`	The data type to assign to time series for	Use the value column names for

	each of the value columns (or specify one value to apply to all columns).	the data types.
Interval	The interval for the time series. Only one interval is recognized for all the time series in the file. Interval choices are provided when editing the command. If it is possible that the date/times are not evenly spaced, then use the IRREGULAR interval.	None – must be specified.
Scenario	The scenario to assign to time series for each of the value columns (or specify one value to apply to all columns).	No scenario will be assigned.
Units	The data units to assign to time series for each of the value columns (or specify one value to apply to all columns).	No units will be assigned.
Missing	Strings that indicate missing data in the file (e.g., "m").	Interpret empty column values as missing data.
Alias	The alias to assign to time series, as a literal string or using the special formatting characters listed by the command editor. The alias is a short identifier used by other commands to locate time series for processing.	No alias will be assigned.

A sample command file to retrieve real-time time-series data from the State of Colorado's website and read the data is as follows:

```
WebGet(URI="http://www.dwr.state.co.us/SurfaceWater/data/export_tabular.aspx?
  IDADATUNCO&MTYPEGAGE_HT,DISCHRG&INTERVAL1&START10/1/06&END10/6/06",
  LocalFile="Data\ Data\CO-DWR-ADATUNCO-tab.txt ")
ReadDelimitedFile(InputFile="Data\CO-DWR-ADATUNCO-tab.txt",
  Delimiter="\t",ColumnNames="ID,
  DateTime,GAGE_HT,DISCHRG",
  DateTimeColumn="DateTime",ValueColumn="GAGE_HT,DISCHRG",
  SkipRowsAfterComments="1",LocationID="ADATUNCO",
  Provider="DWR",DataType="GAGE_HT,DISCHRG",Interval=15Minute,
  Units="ft,cfs",Alias="%L%T")
```