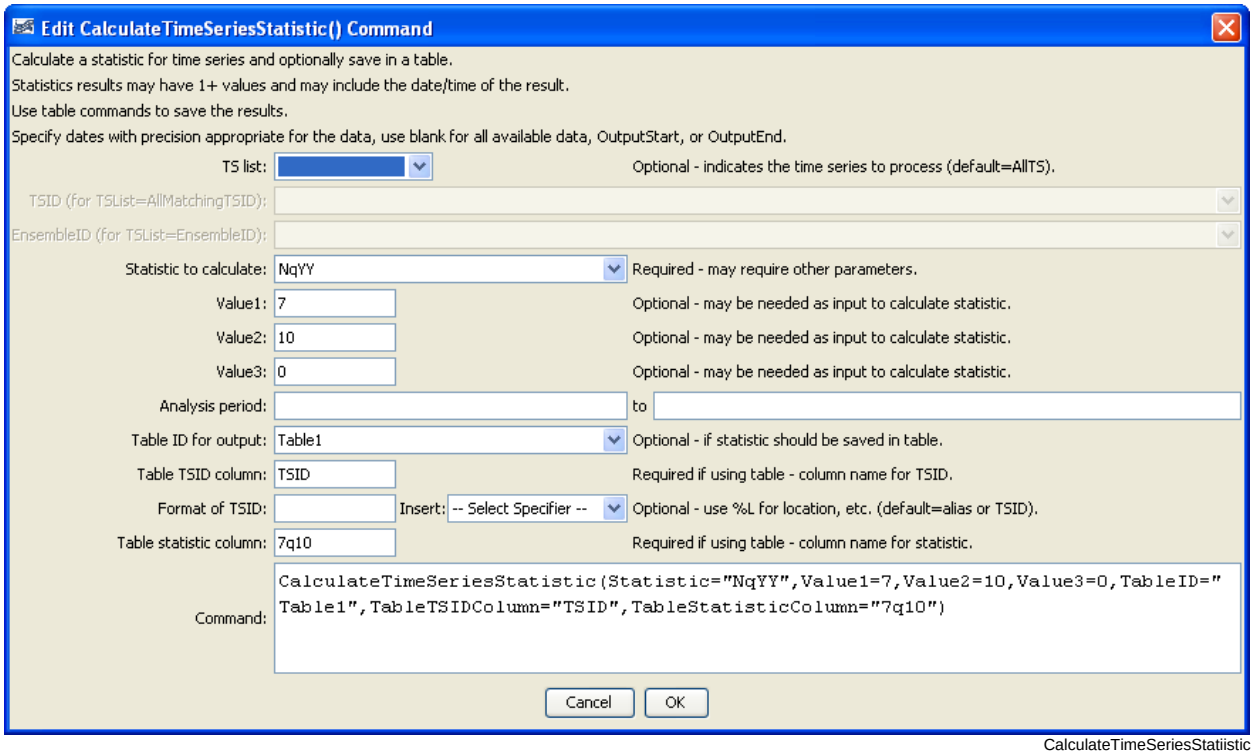
Command Reference: CalculateTimeSeriesStatistic()

Calculate time series statistic

Version 09.08.01, 2010-09-15

The `CalculateTimeSeriesStatistic()` command calculates a statistic for a time series (typically a single value) and optionally adds the result to a table (see the `NewTable()` command). Multiple time series can be processed. The sample from each time series consists of data values for the full period or a shorter period if specified for the command. Missing values are typically ignored unless significant for the statistic (e.g., `Statistic=MissingCount`).

The following dialog is used to edit the command and illustrates the command syntax. Most statistics do not require additional input; however, those that do utilize the `Value*` parameters to specify additional information.



CalculateTimeSeriesStatiistic

CalculateTimeSeriesStatistic() Command Editor

The command syntax is as follows:

```
CalculateTimeSeriesStatistic(Parameter=Value,…)
```

Command Parameters

Parameter	Description	Default
TSList	Indicates the list of time series to be processed, one of: • AllMatchingTSID – all time series that match the TSID (single TSID or TSID with wildcards). • AllTS – all time series before the command. • EnsembleID – all time series in the ensemble. • FirstMatchingTSID – the first time series that matches the TSID (single TSID or TSID with wildcards). • LastMatchingTSID – the last time series that matches the TSID (single TSID or TSID with wildcards). • SelectedTS – the time series selected with the SelectTimeSeries() command.	AllTS
TSID	The time series identifier or alias for the time series to be processed, using the * wildcard character to match multiple time series.	Required if TSList=*TSID.
EnsembleID	The ensemble to be processed, if processing an ensemble.	Required if TSList=EnsembleID.
Statistic	Statistic to compute, one of the following: • Count – number of data values total, including missing and non-missing • DeficitMax – the maximum deficit value (where deficit is mean minus value) • DeficitMean – the mean deficit value (where deficit is mean minus value) • DeficitMin – the minimum deficit value (where deficit is mean minus value) • DeficitSeqLengthMax – the maximum number of sequential intervals where each value is less than the mean (for example maximum drought length) • DeficitSeqLengthMean – the mean number of sequential intervals where each value is less than the mean (for example mean drought length) • DeficitSeqLengthMin – the minimum number of sequential intervals where each value is less than the mean (for example minimum drought length) • DeficitSeqMin – the maximum sum of sequential values where each value is less than	None – must be specified.

Parameter	Description	Default
	the mean (for example maximum drought water volume) • `DeficitSeqMean` – the mean of the sum of sequential values where each value is less than the mean (for example mean drought water volume) • `DeficitSeqMin` – the minimum sum of sequential values where each value is less than the mean (for example minimum drought water volume) • `Lag-1AutoCorrelation` – the autocorrelation between values and the those that follow in the next time step, given by: $r_k = \frac{\sum_{i=1}^{N-k}(Y_i - Y_{mean})(Y_{i+k} - Y_{mean})}{\sum_{i=1}^{N}(Y_i - Y_{mean})^2}$ • `Last` – last non-missing value • `Max` – maximum value • `Mean` – mean value • `Min` – minimum value • `MissingCount` – number of missing values • `MissingPercent` – percent of values that are missing • `NonmissingCount` – number of non-missing values • `NonmissingPercent` – percent of values that are not missing • `NqYY` – restricted to daily data and typically used to analyze return interval of low flows, requires values of `N`, `YY`, and number of missing allowed to be specified with `Value` parameters (see **Statistic Details** table below) • `Skew` – skew coefficient, as follows: $Cs = \frac{N\sum_{i=1}^{N}(Y_i - Y_{mean})^3}{(n-1)(n-2)s^3}$ where s = standard deviation • `StdDev` – standard deviation • `SurplusMin` – the maximum surplus value (where surplus is value minus mean) • `SurplusMean` – the mean surplus value (where surplus is value minus mean) • `SurplusMin` – the minimum surplus value (where surplus is value minus mean) • `SurplusSeqLengthMax` – the maximum number of sequential intervals where each value is greater than the mean (for example maximum water surplus length) • `SurplusSeqLengthMean` – the mean number of sequential intervals where each value	

Parameter	Description	Default
	is greater than the mean (for example mean water surplus length) • `SurplusSeqLengthMin` – the minimum number of sequential intervals where each value is greater than the mean (for example minimum water surplus length) • `SurplusSeqMin` – the maximum sum of sequential values where each value is greater than the mean (for example maximum water surplus volume) • `SurplusSeqMean` – the mean of the sum of sequential values where each value is greater than the mean (for example mean water surplus volume) • `SurplusSeqMin` – the minimum sum of sequential values where each value is greater than the mean (for example minimum water surplus volume) • `Variance` – variance	
`Value1`	Input data required by the statistic. Currently the dialog does not check the value for correctness – it is checked when the statistic is computed.	See **Statistic Details** table below.
`Value2`	Input data required by the statistic. Currently the dialog does not check the value for correctness – it is checked when the statistic is computed.	See **Statistic Details** table below.
`Value3`	Input data required by the statistic. Currently the dialog does not check the value for correctness – it is checked when the statistic is computed.	See **Statistic Details** table below.
`AnalysisStart`	The date/time to start analyzing data.	Full period is analyzed.
`AnalysisEnd`	The date/time to end analyzing data.	Full period is analyzed.
`TableID`	Identifier for table that receives the statistic.	Optional – table output is not required.
`TableTSIDColumn`	Table column name that is used to look up the time series. If a matching TSID is not found, a row will be added to the table. If a TSID is found, the statistic cell value for the time series is modified.	Optional – table output is not required.
`TableTSIDFormat`	The specification to format the time series identifier to insert into the TSID column. Use the format choices and other characters to define a unique identifier.	Time series alias if available, or the time series identifier.
`TableStatistic Column`	Table column name to receive the statistic value. If not found in the table, a new column is added automatically.	Optional – table output is not required.

The following table provides additional information about specific statistics, in particular to describe how the statistic is computed and whether additional input needs to be provided with `Value` command parameters.

Statistic Details

Statistic	Description	Required Values
`NqYY`	This statistic is typically used to evaluate the return period of low flows and is implemented only for daily data. The `N` indicates the number of daily values to be averaged and `YY` indicates the return interval. For example, `7q10` indicates the flow corresponding to the 10-year recurrence interval for minimum average daily flow (for 7 days) in a year. This statistic is computed as follows, using 7q10 as an example: 1. Determine the number of years to be analyzed (from analysis period command parameters or time series data). 2. For each year, loop through each day from January 1 to December 31. Compute an average flow by averaging 7 days, in this case with 3 values on each side of the current day and including the current day. If at the end of the year, use 3 values from adjoining years. The number of missing data allowed is controlled by the `Value3` command parameter. 3. For the year, save the minimum 7-day average. 4. Utilize the minimum values for all years, with log-Pearson Type III distribution, to determine the value for the 10-year recurrence interval. See http://pubs.usgs.gov/sir/2008/5126/section3.html for a description of NqYY and "Hydrology for Engineers, 3rd Edition," Linsley, Kohler, Paulhus for a description of log-Pearson Type III distribution.	`Value1` – specify the number of daily values to be averaged. Currently this must be an odd number to allow bracketing the current day. `Value2` – specify the return interval (e.g., 10). `Value3` – specify the number of missing values allowed in the average (e.g., 0 for most rigorous analysis). It may be useful to set this value if, for example, a single daily value is available in the time series, for example entered on the first day of the month.
All other statistics	Described above.	No additional input values are needed.

The following example illustrates how to use the command to compute the 7q10 statistic for daily flow:

```
TS linsley = ReadDateValue(InputFile="Data\linsley.dv")
  NewTable(TableID="Table1",Columns="TSID,string;7q10,double")
CalculateTimeSeriesStatistic(Statistic="NqYY",Value1=7,Value2=10,Value3=6,
  TableID="Table1",TableTSIDColumn="TSID",TableStatisticColumn="7q10")
WriteTableToDelimitedFile(TableID="Table1",
  OutputFile="Results/Test_CalculateTimeSeriesStatistic_7q10_linsley_out.csv")
```

This page is intentionally blank.