A Vensim® Module to Calculate
Summary Statistics for Historical Fit

ROGELIO OLIVA
Sloan School of Management
Massachusetts Institute of Technology
30 Memorial Drive, E60-355
Cambridge, MA 02142 USA
Phone: (1) 617/253-0834   Fax: (1) 617/252-1998
e-mail: roliva@mit.edu

D-4584

Introduction

Summary statistics of historical fit (Sterman, 1984) have become a standard validity test for system dynamics models. These statistics, however, are also useful to analyze and diagnose the results of partial model calibrations. Partial model calibration normally requires an iterative process of structure formulation, parameter estimation, analysis of fit and residuals, and model re-formulation. The process is repeated until arriving to a structure that adheres to current knowledge about the real system and that is capable of explaining the observed behavior.

The module described here was developed to facilitate an intensive partial model calibration process and full system tests of a model (Oliva, 1996). The module not only calculates the summary statistics for historical fit (Theil inequality statistics) as described by Sterman (1984), but it also generates a series of graphs and tables to facilitate the analysis and diagnosis of fit and residuals.

The purpose of this document is to explain how to use the module and describe the module's output. In no way does this document attempt to teach how to use or interpret the output of the module. To learn about interpretation of the Theil inequality statistics for system dynamics models see Sterman (1984); for econometric estimation and its assumptions see Kennedy (1992); for a general description of time series analysis see Hamilton (1994).

The module takes advantage of Vensim's capability of having more than one model open and was designed to work as a stand-alone module that does not need to be coded into the existing model. This design choice was driven in part by the difficulty of accessing intermediate results in Vensim® macros and the assumption that summary statistics are nor required from every model run. A separate module to analyze goodness of fit eliminates the need for additional model structure that clutters the model and makes simulations slower without loosing the flexibility to use the same formulations over a range of different situations.

The full code for the stand-alone module is listed in two files in the Appendix –

THEIL_MODULE.mdl and **THEIL_MODULE.vgd**. Both files can be used by the Windows or

Macintosh versions of Vensim®. An electronic version of these files can be obtained from the

author at roliva@mit.edu.

How to use the Module

The following screen comes up after opening the **THEIL_MODULE.mdl** file with Vensim® – if

Vensim® is in the text mode use the **View>As Sketch** menu to change to the sketch mode.
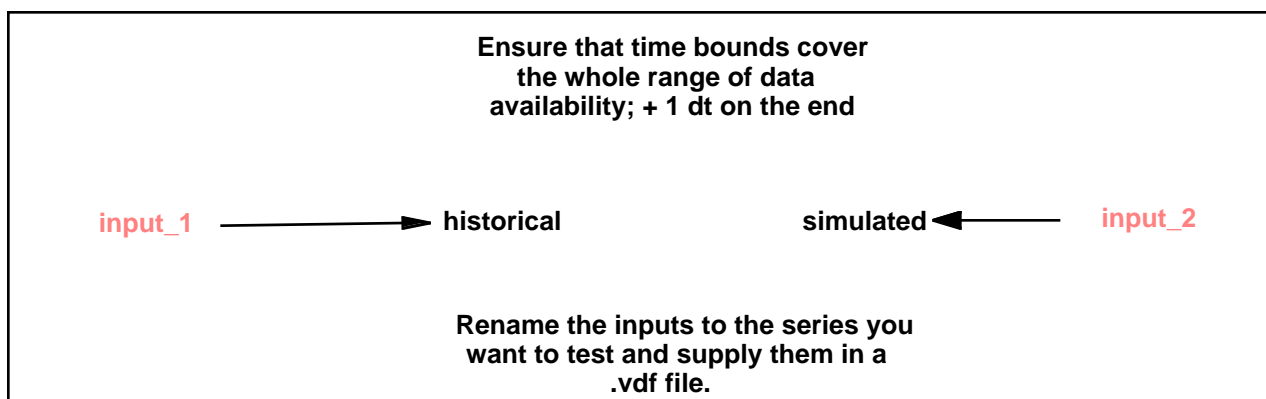
The top message is a remainder to adjust the time bounds of the module to cover the period

where historical data is available and to include an extra simulation step (dt) at the end of the

data availability[1]. The time bounds can be changed using the **Datasets>Time Bounds ...**

menu option. For example, if historical data is available on a weekly basis from week 53 to week

104 the time bounds should be set with to **INITIAL TIME = 53** and **FINAL TIME = 105**. The

TIME STEP for the simulation should be set to equal the minimum interval between two

historical data points.

If the simulated data is available with a higher frequency than the historical data (or over a larger

time horizon), it is possible to reduce the **TIME STEP** (or expand the time bounds) to ensure that

---

[1] The extra dt is required to ensure that all auxiliaries reflect the changes introduced to stocks after the last data point has been read.

the additional simulated data is displayed in the Data Series graph (see Module Output). Time bounds with a wider range than the historical data or a **TIME STEP** smaller than the frequency of the historical data will increase simulation time of the module but will not affect the computational results.

The second message in the opening view indicates how the module is expecting data. The module takes advantage of Vensim's capability to use exogenous data within a model. The assumption is made that the historical and simulated data series are available in a Vensim Data File (`.vdf`) – the simulation output of the model to be tested. If the historical data is not used in the model's simulation it is possible to import the historical data series into the Theil module through a second `.vdf` file – see Chapter 6 Vensim Reference Manual for a description on how to prepare and use data in Vensim models.

Continuing with the above example, lets suppose that the historical data series has a label **"tt_d"** and the simulated series has a label **"time per task,"** and that both series are available in the `tt_opt.vdf` file. To rename the input variables, click on the desired input after having selected the **Variable Tool** from the tool bar and type the new desired name. After going through these changes, the module view for the example looks like this:
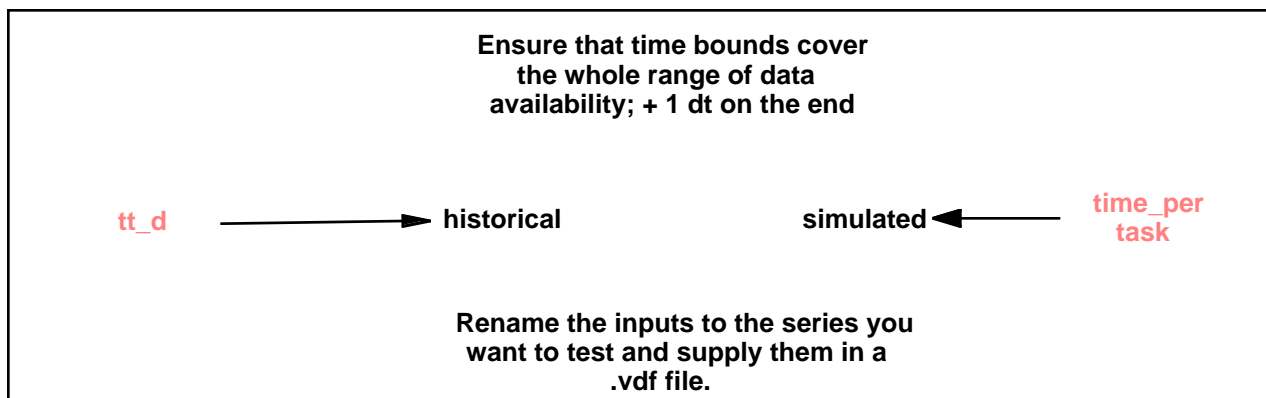
To run the module select the **Datasets>Simulate Model ...** menu. In the **Simulation Control** window specify the file where the time series are available by clicking the **Data ...** button or typing the name of the `.vdf` file directly into the data box. In case the data series are

in different files, specify both files in the data box. This is how the **Simulation Control** window looks after specifying the data file for the example:

Click in the **OK** button to run the module. If you expect to be re-testing these particular series several times – because of structural changes in the model or re-calibration of parameters – it might be a good idea to save the formatted module with a special name to avoid having to specify time bounds and variable names again.

Module Output

This section describes the different tables and graphs that are generated once the module has been simulated with the appropriate changes. Most of the output described on this section is available through customized graphs defined in the **THEIL_MODULE.vgd** file (see Appendix), some graphs, however, are available directly from the Vensim® tool set. To load the **THEIL_MODULE.vgd** graph set use the **Graphs>Load Graph Set** menu when the module is the active model in the Vensim® workbench.

The module is designed to work with only one dataset at a time. Use the **Datasets>Dataset Control** menu to ensure that the proper dataset is being analyzed.

The first report (Summary Statistics) contains the numerical results from the calculations defined by Sterman (1984). The rest of the graphs and reports are included to support the interpretation and diagnosis of the inequality statistics and residuals, and the identification of potential improvements to structural formulations.

Summary Statistics (Report)

```
☐ ⌂ 🖳 ▤ 🖫 ≣ Summary Stat. for Historical Fit ≣ ▲
Summary Statistics for Historical Fit
------------------------------------------------
n =  50.0

R^2                              0.82825
Mean Abs. Percent Error          0.01504
Mean Square Error                0.00039
Root Mean Square Error           0.01979
Bias                             0.00000
Variation                        0.04666
Covariation                      0.95334
------------------------------------------------
```

This report generates the summary statistics for historical fit. By clicking in the 'clipboard' icon the contents of the report can be copied and pasted into a text document. The report provides the following statistics:

n — Number of data points included in the comparison.

Coefficient of Determination (R^2)

$$r^2 = \left(\left(1/n \sum_{t=1}^{n} (S_t A_t) - \overline{SA} \right) \Big/ S_S S_A \right)^2$$

Mean Absolute Percent Error

$$MAPE = 1/n \sum_{t=1}^{n} \left| \frac{S_t - A_t}{A_t} \right|$$

Mean Square Error

$$MSE = 1/n \sum_{t=1}^{n} (S_t - A_t)^2 = (\overline{S} - \overline{A})^2 + (S_S - S_A)^2 + 2(1-r)S_S S_A$$

Root Mean Square Error

$$RMSE = \sqrt{1/n \sum_{t=1}^{n} (S_t - A_t)^2}$$

Bias component of MSE $\quad U^M = (\overline{S} - \overline{A})^2 \big/ MSE$

Variation component of MSE $\quad U^S = (S_S - S_A)^2 \big/ MSE$

Covariation component of MSE $\quad U^C = 2(1-r)S_S S_A / MSE$
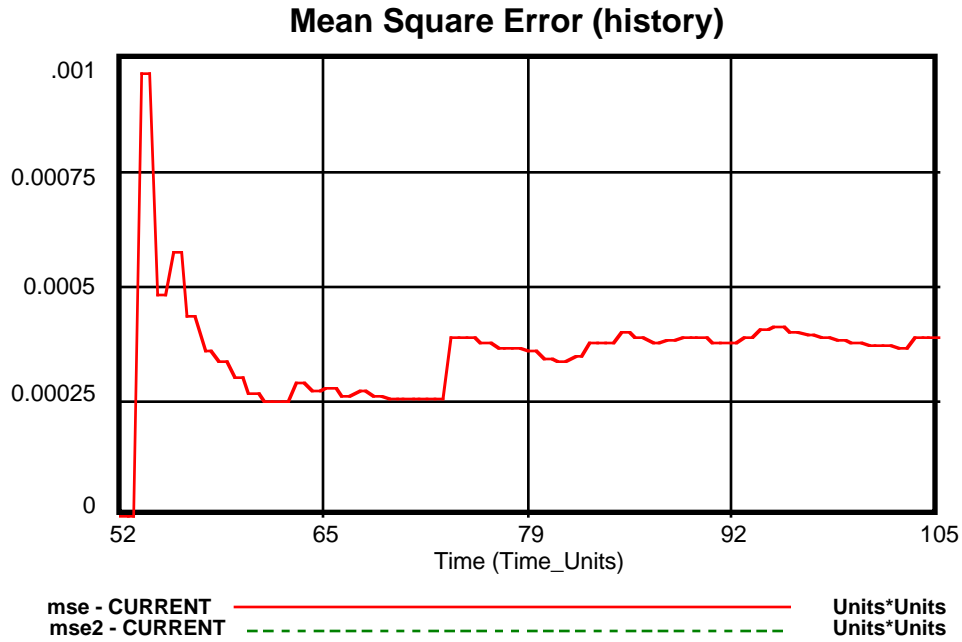
Where:
S_t and A_t represent the value of the simulated and actual (historical) variable at time t,

$\overline{X} = 1/n \sum_{t=1}^{n} X_t$ represents the mean of a time series, and $S_X = \sqrt{1/n \sum_{t=1}^{n} X_t^2 - (\overline{X})^2}$ its standard deviation.

Mean Square Error (history)

Mean Square Error (history)

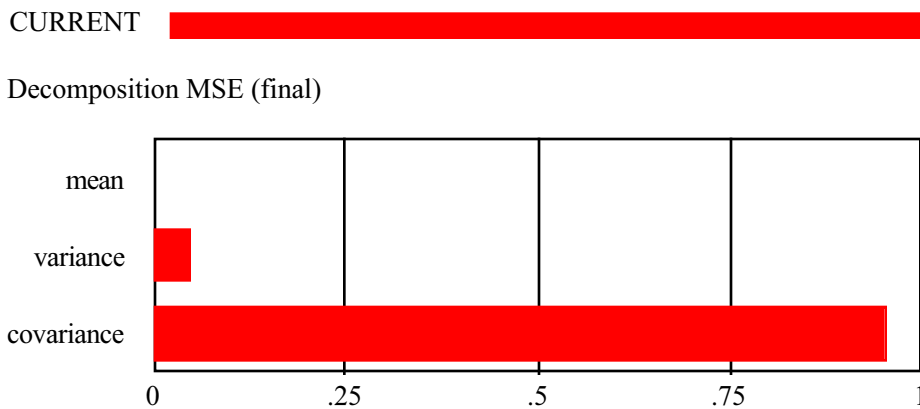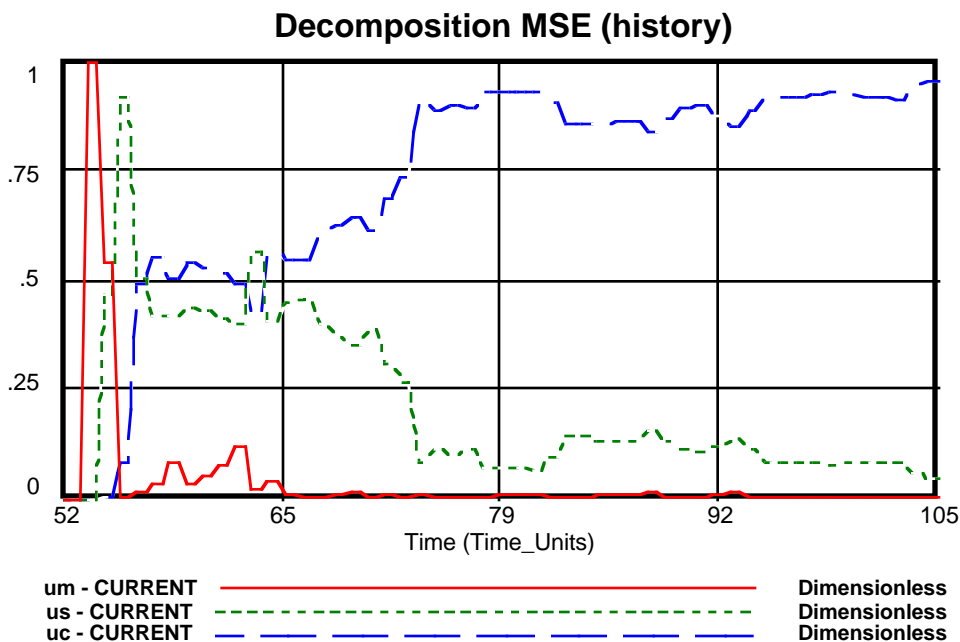|				
mse - CURRENT				Units*Units
mse2 - CURRENT				Units*Units

- Evolution of the Mean Square Error as new points are introduced into the comparison throughout the simulated period.

## Decomposition of Mean Square Error (final)

- Decomposition of the Mean Square Error (MSE) presented graphically. It can be used with the plot of the two data series to interpret the Theil Inequality Statistics (see Figure 1 in Sterman 1984).

Decomposition of Mean Square Error (history)

Decomposition MSE (history)

um - CURRENT ————————	**Dimensionless**
us - CURRENT – – – – – – – –	**Dimensionless**
uc - CURRENT — — — — — —	**Dimensionless**

- Decomposition of the Mean Square Error (MSE) as new points are introduced into the comparison throughout the simulated period. This graph permits to see whether there are changes of the composition of the MSE at different stages of the simulation.

Data Series

Data Series

historical - CURRENT ————————	**Units**
simulated - CURRENT – – – – – – – –	**Units**
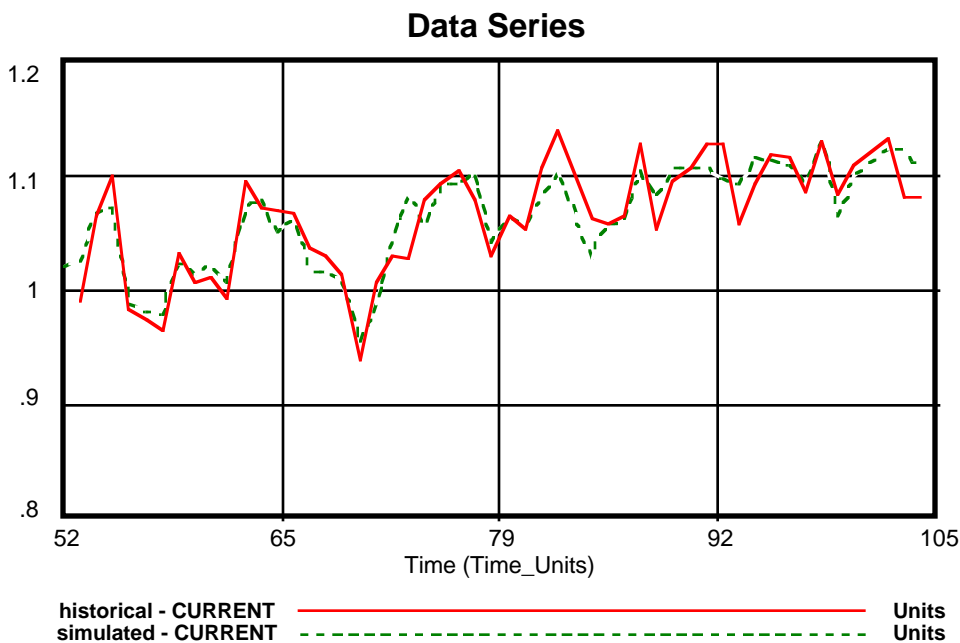
- Graph plotting the actual (historical) and simulated data series. Note that this graph plots the simulated series with the data available within the specified Time Bounds and TIME STEP frequency.

Residuals

Residuals

residuals - CURRENT ———————————————— Units
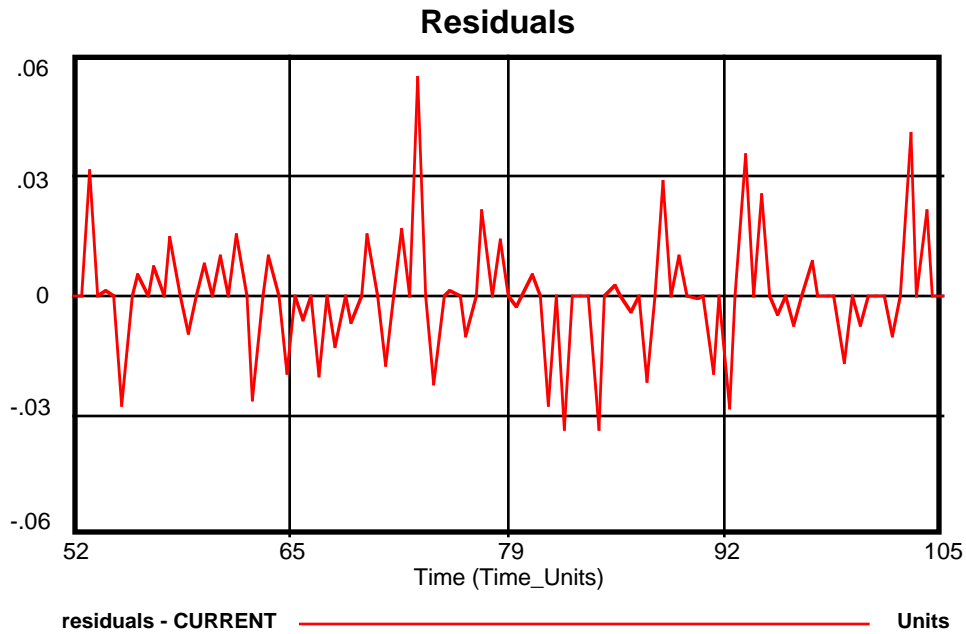
- Plot of residuals over time $res = S_t - A_t$. This graph is useful to detect biases, trends and cyclical components of residuals.

Autocorrelation Spectrum of Residuals

CURRENT ————————————————
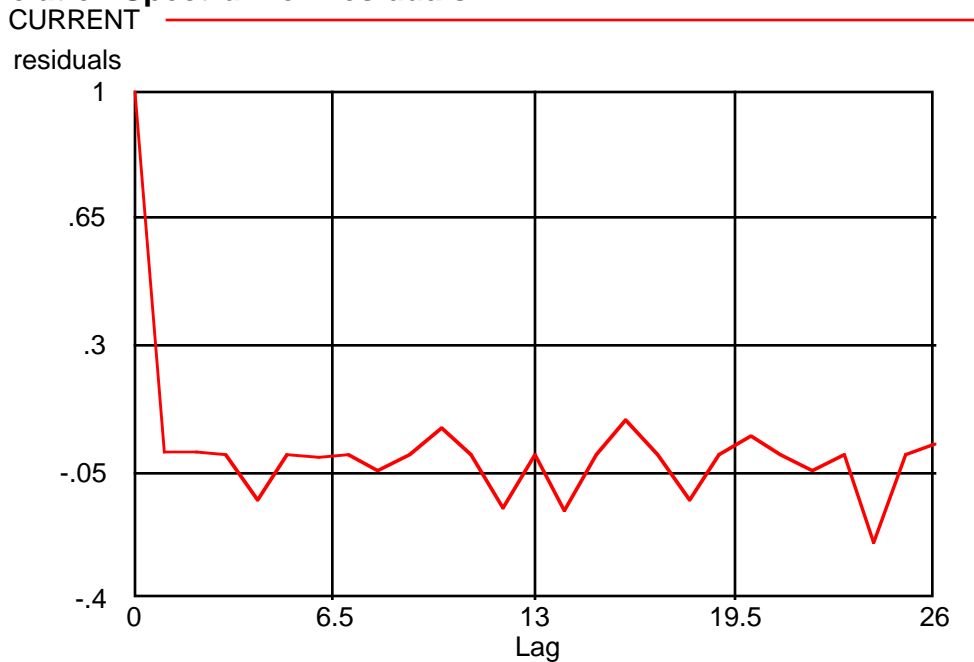
- Autocorrelation spectrum of residuals. This graph is useful to check for cyclical components of the time series and autocorrelated errors. The graph is not available through the customized plots – use the Vensim® **Strip Graph Autocorrelation Option**.

Historical vs. Residuals

Historical vs. Residuals
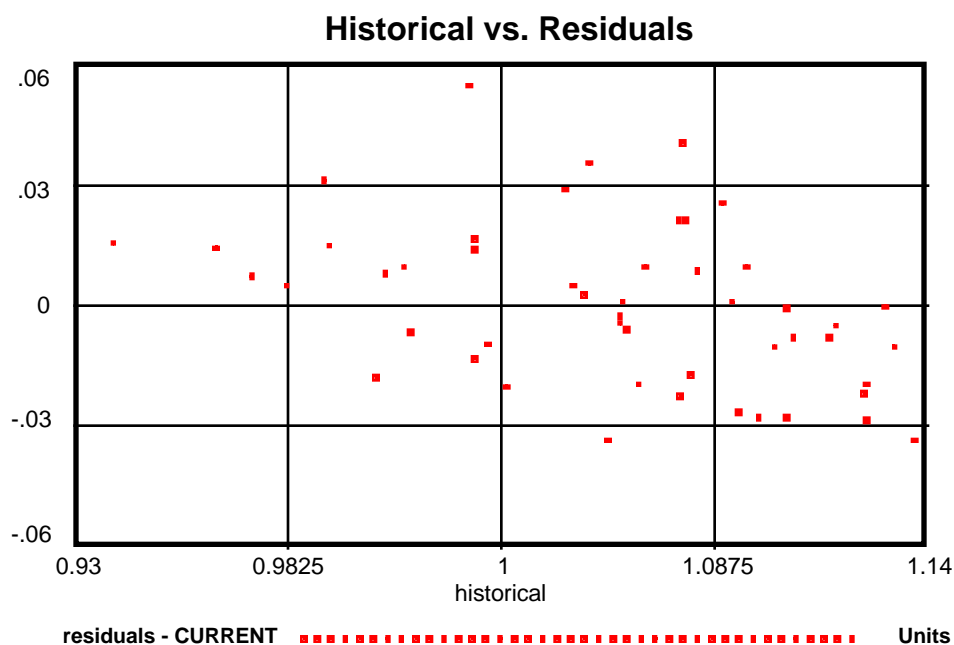
- Scatter plot between the actual values of the variable and the residuals. This graph is useful to check for violations of the equality of variance assumption (heteroskedasticity). It is also a good idea to use the Vensim® **Stats tool** on the residuals to check for mean zero.

References

Hamilton, J.D. 1994. *Time Series Analysis.* Princeton University Press: Princeton, NJ.

Kennedy, P. 1992. *A Guide to Econometrics.* The MIT Press: Cambridge, MA.

Oliva, R. 1996. *A Dynamic Theory of Service Delivery: Implications for Managing Service Quality*. PhD Thesis, Sloan School of Management, Massachusetts Institute of Technology.

Sterman, J.D. 1984. Appropriate Summary Statistics for Evaluating the Historical Fit of System Dynamics Models. *Dynamica*, **10** (Winter), 51-66.

Ventana Systems. 1995. *Vensim 1.62 Reference Manual.* Belmont, MA: Ventana Systems Inc.

Appendix

THEIL_MODULE.mdl

{Module to Calculate Summary Statistics for Historical Fit based on:

Sterman, J.D., (1984) Appropriate Summary Statistics for Evaluating the
Historical Fit of System Dynamics Models. Dynamica, 10 (Winter), 51-66

Rogelio Oliva.
September, 1995}

{Input Series}
input_1
      ~ Units
      ~ Rename this variable to the test data available in the .vdf file
        The series has no value (no '=' sign) to indicate that it will be
        exogenous data. |

input_2
      ~ Units
      ~ Rename this variable to the test data available in the .vdf file
        The series has no value (no '=' sign) to indicate that it will be
        exogenous data. |

historical := input_1
      ~ Units
      ~ Dummy variable to id series |

simulated := input_2
      ~ Units
      ~ Dummy variable to id series |

X := simulated
      ~ Units
      ~ |

Y :RAW: := historical
      ~ Units
      ~ |

{Selection of points through the PICK function}
pick = IF_THEN_ELSE(Y = :NA:, 0, 1)
      ~ Dimensionless
      ~ Flag to id historical value available.
        Takes a value of one for every data point available |

Xi = pick*X
      ~ Units
      ~ Simulated point entering calculations |

Yi = pick*Y
      ~ Units
      ~ Historical point entering calculations |

{Basic Accumulations}
count = INTEG(pick/dt,0)
      ~ Dimensionless
      ~ Counter for # of points |

Sum_Xi = INTEG(Xi/dt,0)
      ~ Units
      ~ Sum of x's (simulated) |

SumX2 = INTEG(Xi*Xi/dt,0)
      ~ Units*Units
      ~ Sum of x^2 (simulated) |

```
Sum_Yi = INTEG(Yi/dt,0)
      ~ Units
      ~ Sum of y's (historical)       |

SumY2 = INTEG(Yi*Yi/dt,0)
      ~ Units*Units
      ~ Sum of y^2 (historical)       |

SumXY = INTEG(Xi*Yi/dt,0)
      ~ Units*Units
      ~ Sum of x*y      |


{Sum of Errors}
Sum_XmY2 = INTEG((Xi-Yi)*(Xi-Yi)/dt,0)
        ~ Units*Units
        ~ Sum of Square Errors (x-y)^2      |

Sum_APE = INTEG(ABS(ZIDZ(Xi-Yi,Yi))/dt,0)
        ~ Dimensionless
        ~ Sum of Absolute Percent Errors |(x-y)/y|      |

Sum_SPE = INTEG((ZIDZ(Xi-Yi,Yi)*ZIDZ(Xi-Yi,Yi))/dt,0)
        ~ Dimensionless
        ~ Sum of Square Percent Errors ((x-y)/y)^2      |


{Calculation of Mean and Standard Deviation}
M_X = ZIDZ(Sum_Xi,count)
      ~ Units
      ~ Mean of x (sum x)/n           |

M_Y = ZIDZ(Sum_Yi,count)
      ~ Units
      ~ Mean of y (sum y)/n           |

MX2 = ZIDZ(SumX2,count)
      ~ Units*Units
      ~ Mean of x^2 (sum x^2)/n       |

MY2 = ZIDZ(SumY2,count)
      ~ Units*Units
      ~ Mean of y^2 (sum y^2)/n       |

Mxy = ZIDZ(SumXY,count)
      ~ Units*Units
      ~ Mean of x*y (sum x*y)/n       |

Sx = SQRT(MX2-(M_X*M_X))
      ~ Units
      ~ Standard Deviation of x. Calculated using the 'hand computation' formula
        to calculate the standard deviation without prior knowledge of the mean.
        Sterman (1984), pg. 64      |

Sy = SQRT(MY2-(M_Y*M_Y))
      ~ Units
      ~ Standard Deviation of y. Calculated using the 'hand computation' formula
        to calculate the standard deviation without prior knowledge of the mean.
        Sterman (1984), pg. 64      |

r = ZIDZ(Mxy-(M_X*M_Y),Sx*Sy)
      ~ Dimensionless
      ~ Correlation coefficient. Calculated through the 'hand computation'.
        Sterman (1984) pg. 63      |


{Decomposition of the Mean Square Error}
dif_mea = (M_X-M_Y)*(M_X-M_Y)
      ~ Units*Units
      ~ Difference of Means (bias)      |
```

12   *© R. Oliva*

```
dif_var = (Sx-Sy)*(Sx-Sy)
      ~ Units*Units
      ~ Difference of variances       |

dif_cov = 2*Sx*Sy*(1-r)
      ~ Units*Units
      ~ Difference of covariances      |

mse = dif_mea + dif_var + dif_cov
      ~ Units*Units
      ~ Mean Square Error. The addition of the three components  |


{Summary Statistics}
RMSPE = SQRT(ZIDZ(Sum_SPE,count))
      ~ Units
      ~ Root Mean Square Percent Error
      ~:SUPPLEMENTARY       |

mape = ZIDZ(Sum_APE,count)
      ~ Dimensionless
      ~ Mean Absolute Percent Error
      ~:SUPPLEMENTARY       |

rmse = SQRT(mse)
      ~ Units
      ~ Root Mean Square Error
      ~:SUPPLEMENTARY       |

um = ZIDZ(dif_mea,mse)
      ~ Dimensionless
      ~ Bias inequality proportion
      ~:SUPPLEMENTARY       |

us = ZIDZ(dif_var,mse)
      ~ Dimensionless
      ~ Variance inequality proportion
      ~:SUPPLEMENTARY       |

uc = ZIDZ(dif_cov,mse)
      ~ Dimensionless
      ~ Covariance inequality proportion
      ~:SUPPLEMENTARY       |

r2 = r*r
      ~ Dimensionless
      ~ Correlation coefficient squared
      ~:SUPPLEMENTARY       |

dt = TIME_STEP
      ~ Time_Units
      ~         |

residuals = Xi-Yi
      ~ Units
      ~ Errors       |


********************************************************
      .Control
********************************************************~
Simulation Control Parameters
      |

FINAL_TIME  = 55
      ~ Time_Units
      ~ The final time for the simulation.          |

INITIAL_TIME  = 0
      ~ Time_Units
      ~ The initial time for the simulation.       |
```

```
SAVEPER  =
        TIME_STEP
     ~ Time_Units
     ~ The frequency with which output is stored.  |

TIME_STEP  = 1
     ~ Time_Units
     ~ The time step for the simulation.          |

\\\---/// Sketch information - do not modify anything except names
V160  Do not put anything below this section - it will be ignored
*View 1
$Helvetica|14|B|0-0-0|0-0-0|0-0-0
|0,"
Rename the inputs to the series you want to test and supply them in a .vdf file.
,429,227,130,23,8,3,0,0,0,-1--1--1,-1--1--1
|1,"
Ensure that time bounds cover the whole range of data availability; + 1 dt on the end
,426,89,115,31,8,3,0,0,0,-1--1--1,-1--1--1
|2,historical,351,161,35,8,0,0,0,0,-1,0,-1--1--1,-1--1--1
|3,simulated,531,161,37,8,0,0,0,0,-1,0,-1--1--1,-1--1--1
|4,input 1,171,162,40,20,8,0,0,1,0,0,255-128-128,-1--1--1
|5,input 2,673,159,40,20,8,0,0,1,0,0,255-128-128,-1--1--1
>4,2,0,0,0,0,0,-1--1--1,1|(256,161)|
>5,3,0,0,0,0,0,-1--1--1,1|(607,159)|
```

```
THEIL_MODULE.vgd

:REPORT Summary Statistics
:TITLE Summary Stat. for Historical Fit
:LOCATION 100,100
:SIZE 350,140
:FONT Monaco|9||0-0-0
Summary Statistics for Historical Fit
-------------------------------------
n = \/

R^2                     \/
Mean Abs. Percent Error \/
Mean Square Error       \/
Root Mean Square Error  \/
Bias                    \/
Variation               \/
Covariation             \/
-------------------------------------
:VAR count %5.1f
:VAR R2 %12.5f
:VAR MAPE %12.5f
:VAR MSE %12.5f
:VAR RMSE %12.5f
:VAR um %12.5f
:VAR us %12.5f
:VAR uc %12.5f
:END-OF-REPORT

:GRAPH MSE_(history)
:TITLE Mean Square Error (history)
:SCALE
:VAR mse
:VAR mse2

:BAR-GRAPH Decomposition_MSE_(final)
:TITLE Decomposition MSE (final)
:Y-MIN 0
:Y-MAX 1
:HORIZONTAL
:VAR um|mean
:VAR us|variance
:VAR uc|covariance

:GRAPH Decomposition_MSE_(history)
:TITLE Decomposition MSE (history)
:SCALE
:VAR um
:Y-MIN 0
:Y-MAX 1
:VAR us
:VAR uc

:GRAPH Data_Series
:TITLE Data Series
:SCALE
:VAR historical
:VAR simulated

:GRAPH Residuals
:TITLE Residuals
:SCALE
:VAR residuals

:GRAPH HISTORICAL_VS._RESIDUAS
:TITLE  Historical vs. Residuals
:X-AXIS historical
:X-LABEL historical
:DOTS
:SCALE
:VAR residuals
```