

User's Manual

EquCheck

Shamik Bhat, Sahitya Talachutla, and Hai Lin

Chemistry Department, University of Colorado Denver, Denver, CO 80217-3364

Program version: 1.0

Program version date: Jan 2, 2019

Manual date: Jan 2, 2019

Copyright: 2019

Abstract: EquCheck is a program for automated checks on whether a molecular dynamics system is in equilibrium with respect to a given variable. The input data are a time-dependent fluctuating signal whose average is a thermodynamic quantity, and they should be given as a time series with equally-spaced time intervals. The implemented algorithms, which are described in the paper by Schiferl and Wallace (*J. Chem. Phys.* **1985**, 83, 5203-5209), include four statistical tests on the coarse-grained (CG) input data: Mann-Kendall test for lack of trend in the CG data mean, Mann-Kendall test for lack of trend in the CG data variance, Shape test for normality of the CG data mean, and one-tailed Von Neumann test for lack of positive correlation in the CG data mean. The four tests were carried out automatically in a successive manner until all tests are passed or all CG samples have run out, whichever comes first. Six test runs are included.

Users are encouraged to send their questions/comments to:

Hai Lin (hai.lin@ucdenver.edu)

Questions about licenses, distribution, and so forth may be sent to:

Hai Lin (hai.lin@ucdenver.edu)

Publications resulting from using the EquCheck program should cite the program. An example in the ACS style is as follows:

Bhat, S.; Talachutla, S.; Lin, H. EquCheck version 1.0, University of Colorado Denver, Denver, 2018.

No guarantee is made that this program is bug-free or suitable for specific applications, and no liability is accepted for any limitations in the mathematical methods and algorithms used within the program. We try our best to provide helps to users, when possible and when time permits; however, no consulting or maintenance services are guaranteed or implied.

The program is available free of charge; however, it should not be redistributed outside of a research group.

Contents

1.	Introduction	4
2.	Theory	5
2.1.	Coarse-Grained Time Series.....	5
2.2.	Mann-Kendall Test for Trend.....	6
2.3.	Shapiro-Wilk (W) Test for Normality	7
2.4.	Shape Test for Normality	8
2.5.	Von Neumann Test for Serial Correlation.....	9
3.	Implementation.....	10
4.	Installation, Input, and Output.....	11
4.1.	Installation	11
4.2.	Input.....	11
4.3.	Output	14
5.	Interpreting Results	18
6.	Test Runs.....	20
6.1.	Test 1: Mikias.dat	20
6.2.	Test 2: Unequ.dat.....	21
6.3.	Test 3: Rand().dat	22
6.4.	Test 4: Rand() + at	23
6.5.	Test 5: Rand() + $\sin(t)$	24
6.6.	Test 5: Rand() + $\ln(t)$	26
7.	Bibliography.....	29
8.	Acknowledgements	29
9.	Appendices	30
	Table A1. Standard normal probabilities for converting one-sided percentiles to z -scores, for both negative and positive z -scores (Standard Normal Probabilities).	31
	Table A2. Values for the conversion of confidence interval to t -score, given degrees of freedom (t -table).	33
	Table A3. Values for a coefficient for n up to 30 for the Shapiro-Wilk test.....	34
	Table A4. Error (α -level) values for the Shapiro-Wilk test, given the sample size n up to 50.	36

1. Introduction

In classical molecular dynamics (MD) simulations, a model system evolves from a starting geometry and initial velocities of the atoms at $t = 0$, generating time series of system variables. (By classical, we mean that the degrees of freedom of the atoms are treated classically by Newtonian mechanics.) Monitoring the time series of the system variables provides insights into the system evolution, and averages of the variables yield thermodynamic quantities of interest. One example is kinetic energy, whose average is proportional to temperature. Another example is the root-mean-square deviation (RMSD) of the protein backbone atoms with respect to their initial positions, whose average is used to indicate the conformational changes of the protein.

The time average to compute the thermodynamic quantity should be taken when the system is in equilibrium, i.e. after a certain equilibration time t_{eq} . The equilibration time is variable specific, as different variables likely have different equilibration times. For example, the system's temperature usually reaches equilibrium quicker than the RMSD of protein backbone atoms does. It is therefore important that we have a practical procedure to assess if the system has reached equilibrium with respect to a given variable.

Let us assume that an MD simulation has generated a time series of a system variable $X(t)$. According to the quasi-ergodic hypothesis, the mean of $X(t)$ is equal to the ensemble average $\langle X \rangle$. Equilibrium is *defined* to be that $X(t)$ is characterized by a probability distribution $P(X)$ that is time-independent. Our goal is to determine the “starting time” t_s , after which the average can be taken. Apparently, $t_s \geq t_{\text{eq}}$.

The algorithms implemented in this EquCheck program are based on the paper by Schiferl and Wallace (J. Chem. Phys. **1985**, 83, 5203-5209), which proposed four statistical tests on the time series of a given variable $X(t)$. A coarse-grained (CG) strategy of $X(t)$ has been employed, and the statistical tests (5% level of significance recommended) are:

- 1) Mann-Kendall test for lack of trend in the CG data mean,
- 2) Mann-Kendall test for lack of trend in the CG data variance,
- 3) W or Shape test for normality of the CG data mean, and
- 4) one-tailed Von Neumann test for lack of positive correlation in the CG data mean.

The four tests were carried out automatically in a successive manner until all tests are passed or all CG samples have run out, whichever comes first. Six test runs are included.

2. Theory

The theory has been detailed in the paper by Schiferl and Wallace (Schiferl and Wallace, 1985) and the other cited references, and here we only provide a brief description.

2.1. Coarse-Grained Time Series

Let us assume that the MD simulations has generated a time series

$$X_i = X(i\Delta t), \quad i = 1, 2, \dots \quad (2.1.1)$$

Here, Δt is the *time interval* for recording the variable X . For example, if the variable is recorded every 10 time steps, the time interval is 10. This differs slightly from the paper by Schiferl and Wallace, where Δt is the *time step* and the variable is recorded every step. Here, we keep time step to its ordinary meaning, i.e. the time step size for trajectory propagation.

A coarse-grained (CG) strategy is employed, as illustrated in Figure 2.1.1.

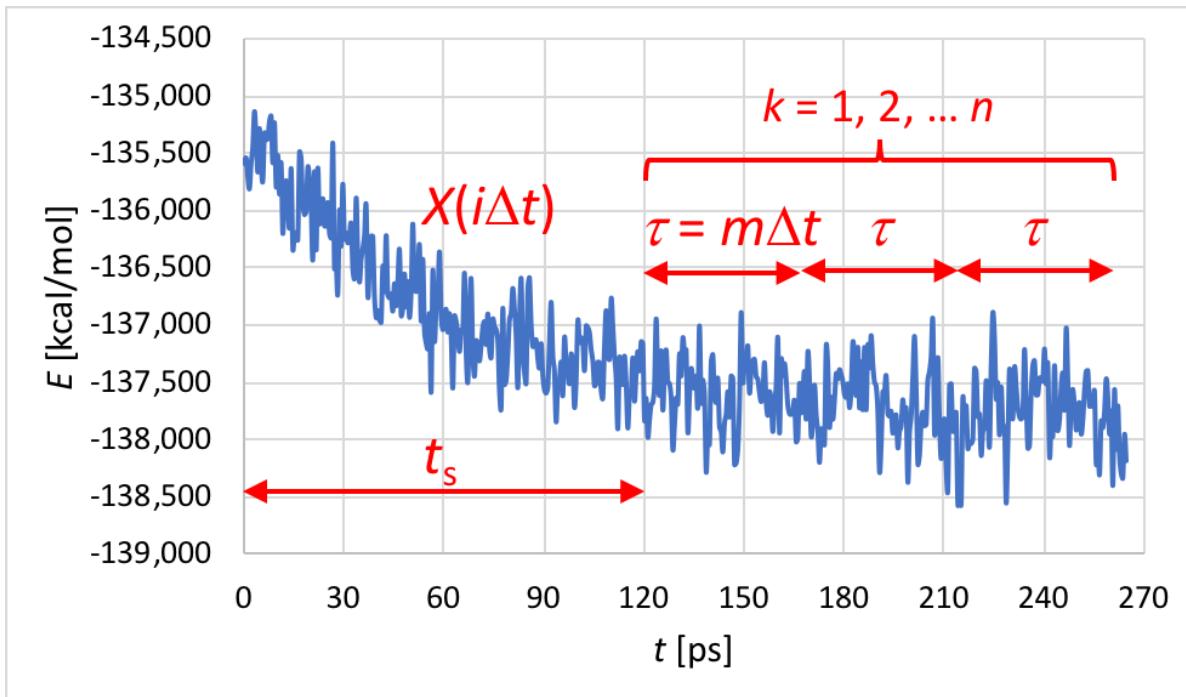


Figure 2.1.1. The coarse-grained strategy exemplified by total energy E recorded every 0.5 ps in a simulation. After the starting time t_s , the series of points $X_i = X(i\Delta t)$ is divided into sequential non-overlapping segments $k, k = 1, 2, \dots, n$. Each segment has m points, and the length is $\tau = m\Delta t$.

Basically, after discarding the first portion of the series from $t = 0$ to t_s , which is considered to be the equilibration stage, one divides the rest of the series points into n sequential non-overlapping segments. Each segment has m points, and the length is $\tau = m\Delta t$. The recommendation is $n \geq 24$ and $\tau \approx t_f$, where

t_f is the characteristic fluctuation time of the variable $X(t)$. The fluctuation time can be estimated by dividing the elapsed time between two data points at t_1 and t_2 by C_{\max} , the number of local maxima counted during this time period:

$$t_f = \frac{t_2 - t_1}{C_{\max}} \quad (2.1.2)$$

The segment is labelled by $k, k = 1, 2, \dots, n$, of which the mean is x_k and the variance is S_k^2 , respectively. These x_k values constitute a sample drawn from a population whose mean is the mean of the raw MD data $\langle X \rangle$. The sample mean \bar{x} and variance s^2 are

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k \quad (2.1.2)$$

$$s^2 = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x})^2 \quad (2.1.3)$$

The equilibrium and the confidence interval of $\langle X \rangle$ are determined based on the statistical analysis of the sample of coarse-grained data. Four tests are carried out, which are described in the rest of this section. If all 4 statistical tests pass, the system will be considered to be equilibrated, and mean variable values will be provided with the calculated confidence interval:

$$\langle X \rangle = \bar{x} \pm a_n \frac{s}{\sqrt{n}} \quad (2.1.4)$$

where a_n is the 0.975 quantile of student's t -distribution with $n - 1$ degrees of freedom. For $n \geq 24$, $a_n \approx 2$. The standard error of the mean, $a_n \frac{s}{\sqrt{n}}$, reveals the error associated with the thermodynamic average that is associated with calculated the sample average with coarse-grained data.

2.2. Mann-Kendall Test for Trend

The **Mann-Kendall** test is a nonparametric test. In other words, the data does not have to be normally distributed for conclusions to be drawn.

The test statistic S is given by

$$S = 2I - \binom{n}{2} \quad (2.2.1)$$

Here, I is the number of distinct increasing pairs of observations (x_j, x_k) , where $x_j < x_k$, $j < k$. For example, suppose that there are 4 segments ($n = 4$) and $x_1 = 1.21, x_2 = 1.19, x_3 = 1.22$, and $x_4 = 1.20$. Among the 6 pairs $(x_1, x_2), (x_1, x_3), (x_1, x_4), (x_2, x_3), (x_2, x_4)$, and (x_3, x_4) , 3 are increasing: (x_1, x_3) , (x_2, x_3) , and (x_2, x_4) , leading to $I = 3$, and $S = 2 \times 3 - \binom{4}{2} = 0$.

For $n \geq 20$, S is approximately normally distributed with mean $\bar{S} = 0$ and standard deviation

$$\sigma_S = \left(\frac{n(n-1)(2n+5)}{18} \right)^{\frac{1}{2}} \quad (2.2.2)$$

For MD simulations, we compute the final test statistic

$$u(S) = S/\sigma_s \quad (2.2.3)$$

which can be compared to widely available Z-score and *t*-score tables to determine whether the result obtained is statistically significant. The final test statistic $u(S)$ should be around 0. If this is true, there is no trend in the data. If not, there is a trend, and the time series fails the Mann-Kendall test. Technically, $u(S)$ should be below the Z-score for the significance α level decided by the user. For example, the $\alpha = 0.05$ level correlates to a Z-score of 1.96. If $-1.96 < u(S) < 1.96$, the time series passes the Mann-Kendall test; otherwise, it fails.

The Mann-Kendall test will be used for both the coarse-grained means and the coarse-grained variances. For the latter, the pairs will be (S_j^2, S_k^2) .

The null hypothesis here is that the sample does not show a trend (in the means or variances) while the alternative hypothesis is that the sample does show a trend (Schiferl and Wallace, 1985). To pass this test, the null hypothesis must be accepted.

2.3. Shapiro-Wilk (*W*) Test for Normality

The **Shapiro-Wilk** test (Shapiro and Wilk, 1965) is a test for normality of the data, which should be used when $n \leq 50$. For $n > 50$, another test, the Shape test, is recommended.

First, we rearrange the observation x_k in ascending order: $x_1' < x_2' < x_3' \dots < x_n'$ and calculate the slope of the probability plot regression line:

$$b = \sum_{k=1}^l a_{n-k+1} (x_{n-k+1}' - x_k'), \text{ where } l = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ \frac{(n-1)}{2} & \text{if } n \text{ is odd} \end{cases} \quad (2.3.1)$$

The test statistic W is given by

$$W = \frac{b^2}{s^2(n-1)} \quad (2.3.2)$$

Here, s^2 is the sample variance of the distribution, and a_{n-k+1} are the coefficients tabulated in (Shapiro and Wilk, 1965).

The null hypothesis here is that the sample is drawn from a normal distribution while the alternative hypothesis is that the population is not normal (Schiferl and Wallace, 1985). To pass this test, the null hypothesis must be accepted. The test statistic W should generally be less than 0.95. If this is true, there is a lack of evidence of non-normality in sample means, and the data pass the test.

2.4. Shape Test for Normality

The **Shape** test is a test for normality for when $n \geq 50$, for which the calculations for the a_{n-k+1} coefficients and the critical values for W become very complicated (Schiferl 1985). The shape test looks at both skewness (how symmetric the distribution is) and kurtosis (how sharp the central peak is).

First, we calculate the 2nd to 4th moment about the mean: $m_j, j = 2, 3, 4$.

$$m_j = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x})^j \quad (2.4.1)$$

Next, we calculate the population skewness g_1 and sample skewness G_1 . Dividing G_1 by the standard error of skewness (SES) gives a Z-test statistic, Z_{g1} , to compare with the critical values. (Brown 2016)

$$g_1 = \frac{m_3}{(m_2)^{3/2}} \quad (2.4.2)$$

$$G_1 = \frac{\sqrt{n(n-1)}}{n-2} g_1 \quad (2.4.3)$$

$$Z_{g1} = \frac{G_1}{\text{SES}} \quad (2.4.4)$$

$$\text{SES} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n-3)}} \quad (2.4.5)$$

Similarly, for kurtosis, we calculate the population kurtosis a_4 , population excess kurtosis g_2 , and sample excess kurtosis G_2 . Dividing G_2 by the standard error of kurtosis (SEK) leads to the test statistic Z_{g2} .

$$a_4 = \frac{m_4}{(m_2)^2} \quad (2.4.6)$$

$$g_2 = a_4 - 3 \quad (2.4.7)$$

$$G_2 = \frac{n-1}{(n-2)(n-3)} [(n+1)g_2 + 6] \quad (2.4.8)$$

$$Z_{g2} = \frac{G_2}{\text{SEK}} \quad (2.4.9)$$

$$\text{SEK} = 2(\text{SES}) \sqrt{\frac{n^2-1}{(n-3)(n+5)}} \quad (2.4.10)$$

If the absolute values of Z_{g1} and Z_{g2} are smaller than the respective critical values, both of which are 2 (~5% significance level), the test passes.

The assumptions for the Shape test are the same as those of the Shapiro-Wilk test. The null hypothesis and the alternative hypothesis are the same as well. If these two test statistics are smaller than the respective critical values, there is a lack of evidence of non-normality in the sample means, and thus the data passes the test for normality.

Note that the skewness G_1 or g_1 tells you how highly skewed your sample is: the bigger the number, the more highly the skew. The test statistic Z_{g1} tells you whether the distribution is probably skewed, but not by how much: the bigger the number, the higher the probability. Similarly, the excess kurtosis G_2 or g_2 tells you how the sharp the peak is compared with the peak in the normal distribution ($\alpha_4 = 3$): the more positive the number, the sharper the peak. The test statistic Z_{g2} tells you whether the whole distribution is probably platykurtic, but not by how much: the bigger the number, the higher the probability.

2.5. Von Neumann Test for Serial Correlation

The **Von Neumann test** is a test for lack of correlation between adjacent observations (the segment means). The test assumes that the data is normally distributed, so the data must pass the W or Shape test before the Von Neumann test is performed. The null hypothesis is that there is no such correlation for the segment averages (Schiferl and Wallace, 1985).

First, we calculate the test statistic r

$$r = \frac{q^2}{s^2} \quad (2.5.1)$$

where s^2 is the sample variance, and q^2 is the mean square successive difference

$$q^2 = \frac{1}{2(n-1)} \sum_{k=1}^{n-1} (x_{k+1} - x_k)^2 \quad (2.5.2)$$

For $n \geq 20$, r is approximately normally distributed with mean of 1 and standard deviation σ_r given by

$$\sigma_r = \left(\frac{1}{n+1} \left(1 + \frac{1}{n-1} \right) \right)^{1/2} \quad (2.5.3)$$

Next, we compute the standard normal deviate $u(r)$

$$u(r) = \frac{r-1}{\sigma_r} \quad (2.5.4)$$

If the absolute value of the standard normal deviate, $|u(r)|$, is smaller than the critical value Z of the significance level, there is a lack of correlation in the segment means. The Z value is 1.96 for a significance α level of 0.05. The corresponding values to the other significance levels can be found in the standard normal distribution at Table A1.

3. Implementation

The EquCheck code is written in Perl 5. As shown in the flow chart (Figure 2), the program carries out the tests in the following order

- 1) Mann-Kendall test for lack of trend in the CG data mean,
- 2) Mann-Kendall test for lack of trend in the CG data variance,
- 3) *W* or Shape test for normality of the CG data mean, and
- 4) one-tailed von Neumann test for lack of positive correlation in the CG data mean.

The tables of the critical values for these tests are quoted from the original literature and are provided in the Appendix of this manual.

The code has been tested with sample data sets. Six tests are included in the distribution of this version of EquCheck.

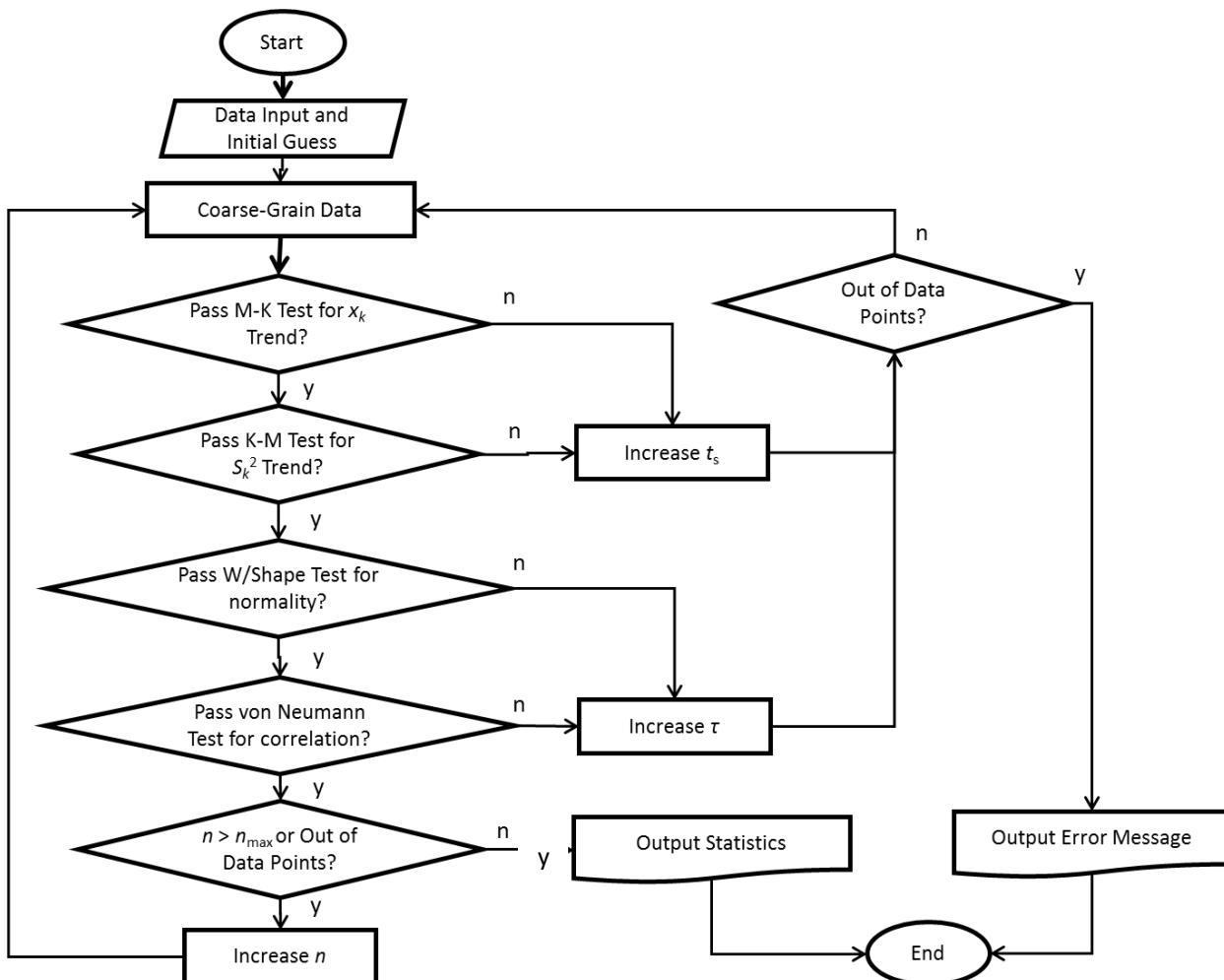


Figure 3.1. Flow chart of EquCheck.

4. Installation, Input, and Output

4.1. Installation

The EquCheck package distribution is provided as a compressed file: EquCheck_x.y.z.tgz, where x , y , and z are version numbers. The current version is 1.0.0.

To install (taking version 1.0.0 as example), under a desired directory (e.g. the user's home directory \sim), unzip the package by typing

```
tar xvfz EquCheck_1.0.0.tgz
```

This will create a directory EquCheck, under which there are 4 subdirectories, bin, doc, test, and testo.

- The subdirectory bin contains the Perl script EquCheck.pl. One should make this path accessible, e.g. by adding the following line to the file .bashrc (for bash shell), .zshrc (for Z shell), or .kshrc (for korn shell):

```
export PATH=$PATH:~/EquCheck/bin
```

or adding the following line to the .cshrc (for C shell) or .tcshrc (for Tenex shell):

```
set path = ($path ~/EquCheck/bin)
```

- The subdirectory doc contains the user's manual EquCheck_UserManual.docx and EquCheck_UserManual.pdf.
- The subdirectory test contains testing data sets: test01.dat to test06.dat.
- The subdirectory test contains sample testing output files obtained by the developers: test01_EquCheck.out to test06_EquCheck.out.

4.2. Input

The EquCheck program runs interactively in a Perl shell environment. To begin, the user provides the information on the data file that contains the recorded time series and other parameters to initialize the data coarse-graining.

The input data file is a text file. Each line is a record, containing two numbers separated by space(s): the first number is number of steps at which the record is taken, and the second number is the value of the time series variable. For example, in test01.dat:

```
0 -135559.0862
500 -135586.1864
1000 -135541.3250
1500 -135710.7772
2000 -135805.9480
...
```

When a user starts the EquCheck program, by typing (all user inputs are in red color in this manual)

`perl EquCheck.pl`

the program displays the following welcome message

Welcome to EquCheck 1.0.0.

This program examines if a time series recorded in MD simulation reaches equilibrium.

Written by Shamik Bhat, Sahit Talachutla, Hai Lin.

Last code change at 12/19/2018.

Please send any comments to Prof. Hai Lin (`hai.lin@ucdenver.edu`).

Next, the program checks and reports the current working directory for user's record

The current working directory of EquCheck.pl is
`/Users/hailinhai/hlin/qmmm/EquCheck/EquCheck/test`

Would you like to change the directory? Answer Y or N (default).

The user can change working directory if desired.

The program then lists the file under the current directory and asks for the name of the data file:

Type the file name within directory!

.

..

`test06.dat`

`test05.dat`

`test04.dat`

`test03.dat`

`test02.dat`

`test01.dat`

`test01.dat`

After typing the answer “`test01.dat`” and hit the “Return” key, the program will open and read the data from the input file. We strongly recommend that the EquCheck program is being run from the same directory where the data file is stored in order to minimize the chance of file access failure. The program first checks the records in the input file for any extra spaces, commas, or other separators and removes them if they are present. The program then parses through the data and inserts each value in a 2×2 array. All the statistical tests will be carried out on this array.

Next, the program asks for sample size information:

What is the sample size (i.e. the number of fragments for data coarse-graining)? This must be 24 (default) or larger.

`25`

Would you like to impose a limit on the sample size? Please answer Y or N. Default is N.

Y

What is the maximal sample size? Default is the current sample size.

25

The sample size is the number of segments n (default is 24), and the maximal sample size is n_{\max} . A large sample size will lead to more powerful von Neumann test and higher confidences in the equilibration and mean value. However, a large sample size may slow down the analysis significantly.

Next, the program outputs the time intervals in time steps between the adjacent data points in the data series and asks for the segment length:

The time interval (in time steps) between adjacent data points is 500.

What is your segment length (fluctuation time) in time steps? For example, if the time series is recorded every 10 time steps, and if the fluctuation time is 200 time steps, each fragment will have $200/10 = 20$ data points.

The default segment length is $2 \times (\text{time interval})$.

500

The initial segment length τ (in the unit of time step) should be set to approximately the same as the estimated fluctuating time Δt_f .

Next, the program asks for the significant alpha level:

What is the alpha level? This is probability of making the wrong decision when the null hypothesis is true. Allowed values are 0.01, 0.02, 0.05 (default), 0.10, and 0.50.

0.05

The alpha level is the significance level, i.e. the probability of rejecting the null hypothesis when the null hypothesis is true.

Next, the program asks for the estimated starting time:

What is the estimated time (in time steps) when the system reaches equilibrium?

Default value is 0 (i.e. the system is already in equilibrium).

0

The start time t_s (in the unit of time step) for equilibration should be estimated through visual inspection of the plot of the time series, or simply set to 0 (default) to initiate the test.

Next, the program asks how detailed the output should be:

Please select the level (1, 2, or 3) of output details? Level 1 is brief, level 2 (default) is modest, and level 3 is very detailed.

2

The program offers 3 level of details for the output: Level 1 reports only the basic statistical information; level 2 also reports changes whenever sample size is changed; and the most verbose level 3 reports all diagnostic information. Level 1 should be sufficient for most applications, while levels 2 and 3 are mostly used by developers for debugging.

4.3. Output

After gathering and verifying the necessary parameters, the program performs the 4 tests, automatically adjusting the stating time t_s , segment length τ , and sample size n (number of segments) when needed:

- If Mann-Kendall tests are failed, stating time t_s will be increased by the time interval Δt before repeating the test.
- If the Mann-Kendall tests are passed, but the W (or Shape) test and/or the von Neumann test fail, the segment length τ will be increased by the time interval Δt before repeating the test.
- If all data points have been used, but there is at least one failed test, the model system is deemed not equilibrated.
- If all tests are passed, and there is no request to increase the sample size, the program terminates with a message that the analysis has been successful and that the model is in equilibrium as determined at the user-specified confidence level.
- If all tests are passed, and there is a request to increase the sample size, the program will increase sample size n by 1 (until n reaching the maximal sample size or the time series reaching the end, whichever comes first) while keeping segment length τ constant and redo the tests. When reaching the maximal sample size or when all data points have been used, the program terminates with a message that the analysis has been successful (with the largest n passing all tests) and that the model is in equilibrium as determined at the user-specified confidence level.

The output is displayed on the screen and saved in an output file `EquCheck.out` under the same directory. Figures 4.2.1 to 4.2.3 show sections of the sample outputs at levels 1 to 3, respectively, with the input parameters above.

```
SUMMARY OF FINAL RESULTS
USER INPUT
Original Sample Size = 24
Sample Size Cap = 25
Time Interval = 500 time units
Segment Length = 5000 time units
Alpha Level = 0.05
The User Estimated Start Time = 0 time units

Your data is equilibrated! RELEVANT STATISTICAL INFORMATION BELOW

The start time is 232 (line number) or 116000 (actual time step)
The sample average is -137648.5565984 with an error of 72.7370101040329
The sample variance is 23337.9187546562
```

The total number of segments is 25
 The length of each segment is 10 data points

THE TEST STATISTICS BELOW

The alpha level you chose is 0.05
 The Z Score used is 1.959964
 The T Score used is 2.069

Mann-Kendall Test for Averages
 The Mann-Kendall Statistic for Averages is -1.91510740263735
 $-1.91510740263735 > -1.959964$, so the test passes.

Mann-Kendall Test for Variance
 The Mann-Kendall Statistic for Variance is 0.887488796344136
 $0.887488796344136 < 1.959964$, so the test passes.

Shapiro-Wilk Test for Normality
 The Test Statistic for Shapiro Wilk is 0.174491361466106
 $0.174491361466106 < 0.916$, so the test passes.

The Von Neumann Test for Trend in Segment Averages
 The Von Neumann Test Statistic is -1.54624890524695
 $-1.54624890524695 > -1.959964$, so the test passes.

Figure 4.2.1. Example of Level 1 output.

NEW RUN
 Sample Size = 24
 Segment Length (divided by Time Interval) = 10 data points
 The Start Time Interval = 239 time intervals
 Alpha Level = 0.05
 The Z Score used is 1.959964
 The T Score used is 2.069

SUMMARY OF FINAL RESULTS
 USER INPUT
 Original Sample Size = 24
 Sample Size Cap = 25
 Time Interval = 500 time units
 Segment Length = 5000 time units
 Alpha Level = 0.05
 The User Estimated Start Time = 0 time units

Your data is equilibrated! RELEVANT STATISTICAL INFORMATION BELOW

The start time is 232 (line number) or 116000 (actual time step)
 The sample average is -137648.5565984 with an error of 72.7370101040329
 The sample variance is 23337.9187546562
 The total number of segments is 25
 The length of each segment is 10 data points

THE TEST STATISTICS BELOW

The alpha level you chose is 0.05
 The Z Score used is 1.959964
 The T Score used is 2.069

Mann-Kendall Test for Averages
 The Mann-Kendall Statistic for Averages is -1.91510740263735
 $-1.91510740263735 > -1.959964$, so the test passes.

```
Mann-Kendall Test for Variance
The Mann-Kendall Statistic for Variance is 0.887488796344136
0.887488796344136 < 1.959964, so the test passes.

Shapiro-Wilk Test for Normality
The Test Statistic for Shapiro Wilk is 0.174491361466106
0.174491361466106 < 0.916, so the test passes.

The Von Neumann Test for Trend in Segment Averages
The Von Neumann Test Statistic is -1.54624890524695
-1.54624890524695 > -1.959964, so the test passes.
```

Figure 4.2.2. Example of Level 2 output.

```
DIAGNOSTIC INFORMATION
The start time step is now 0
The sample size is now 24
The segment length is now 10

DIAGNOSTIC INFORMATION
The start time step is now 1
The sample size is now 24
The segment length is now 10

...
DIAGNOSTIC INFORMATION
The start time step is now 239
The sample size is now 24
The segment length is now 10

NEW RUN
Sample Size = 24
Segment Length (divided by Time Interval) = 10 data points
The Start Time Interval = 239 time intervals
Alpha Level = 0.05
The Z Score used is 1.959964
The T Score used is 2.069

DIAGNOSTIC INFORMATION
The start time step is now 0
The sample size is now 25
The segment length is now 10

DIAGNOSTIC INFORMATION
The start time step is now 1
The sample size is now 25
The segment length is now 10

...
DIAGNOSTIC INFORMATION
The start time step is now 231
The sample size is now 25
The segment length is now 10

SUMMARY OF FINAL RESULTS
USER INPUT
Original Sample Size = 24
```

```
Sample Size Cap = 25
Time Interval = 500 time units
Segment Length = 5000 time units
Alpha Level = 0.05
The User Estimated Start Time = 0 time units

Your data is equilibrated! RELEVANT STATISTICAL INFORMATION BELOW

The start time is 232 (line number) or 116000 (actual time step)
The sample average is -137648.5565984 with an error of 72.7370101040329
The sample variance is 23337.9187546562
The total number of segments is 25
The length of each segment is 10 data points

THE TEST STATISTICS BELOW

The alpha level you chose is 0.05
The Z Score used is 1.959964
The T Score used is 2.069

Mann-Kendall Test for Averages
The Mann-Kendall Statistic for Averages is -1.91510740263735
-1.91510740263735 > -1.959964, so the test passes.

Mann-Kendall Test for Variance
The Mann-Kendall Statistic for Variance is 0.887488796344136
0.887488796344136 < 1.959964, so the test passes.

Shapiro-Wilk Test for Normality
The Test Statistic for Shapiro Wilk is 0.174491361466106
0.174491361466106 < 0.916, so the test passes.

The Von Neumann Test for Trend in Segment Averages
The Von Neumann Test Statistic is -1.54624890524695
-1.54624890524695 > -1.959964, so the test passes.
```

Figure 4.2.3. Example of Level 3 output.

5. Interpreting Results

If, after running out all allowed values of the adjustable parameters (stating time t_s , segment length τ , and sample size n), the time series still cannot pass all four tests, the model system is deemed not equilibrated in terms of the given variable. Failing to pass the Mann-Kendall tests indicates that the time series has a trend throughout the simulations. This implies significant issues in the simulation setup. If the Mann-Kendall tests are passed, but the W (or Shape) test and/or the von Neumann test fail, the simulation should be run for longer time so that segment length can be increased for the next-round test (Schiferl and Wallace, 1985).

If all 4 statistical tests pass, the system will be considered to be equilibrated. Mean variable values will be provided with the calculated confidence interval:

$$\langle X \rangle = \bar{x} \pm a_n \frac{s}{\sqrt{n}} \quad (5.1)$$

As described earlier in this manual, \bar{x} is the sample mean, s^2 is the sample variance, n is the sample size, and a_n is the 0.975 quantile of Student's t -distribution with $n - 1$ degrees of freedom. For $n \geq 24$, $a_n \approx 2$. One can increase the confidences in the equilibration and mean value by increasing the sample size cap and rerun the program.

Here, we especially emphasize the importance of an accurate estimation of the relaxation time, which determines the minimal segment length. The default value has been set to the absolute minimum value of 2, because there is no way for the code developers to foretell the relaxation time. But using this default value is highly discouraged (don't be that lazy, please!), which may lead to the data series falsely passing the tests.

To simplify the calculations in this code, we have set a_n to the t -score value corresponding to 23 degrees of freedom, even when larger sample sizes ($n > 24$) are in use. For example, at the 95% confidence level, we use $a_n = 2.064$ even though more accurate a_n values can be obtained from Table A2. This approximation leads to a slight overestimation of the error bar. For example, at the same 95% confidence level, $a_n = 1.960$ for $n = 1000$, and the error bar is overestimated by only ~5%.

Similarly, the critical value W for the Shapiro-Wilk test corresponds to a sample size $n = 24$, although more accurate larger W values can be obtained from Table A4 for larger sample sizes. The approximation leads to slightly stricter (and safer) test. For example, at the 95% confidence level, $W = 0.984$ for $n = 24$, as compared with $W = 0.988$ for $n = 50$.

Use of more accurate a_n and W values will be considered in the future version of the code.

6. Test Runs

Six test runs are included in the distribution of this version of EquCheck, which are described below.

6.1. Test 1: Mikias.dat

The data in this test were the total energies of a model system obtained from a simulation by a group member (Mikias Negussie), and they are presented in a file named `test01.dat`.

```
0 -135559.0862
500 -135586.1864
1000 -135541.3250
1500 -135710.7772
2000 -135805.9480
...
264500 -138180.9623
265000 -138347.3909
```

Visual inspection of the figure suggests that the series may be equilibrated after some time. Indeed, the time series passed all four tests with the following input parameters without a cap on sample size:

```
Original Sample Size = 24
Time Interval = 500 time units
Segment Length = 5000 time units
Alpha Level = 0.05
The User Estimated Start Time = 0 time units
```

A plot of the time series is given in Fig. 6.1.1.

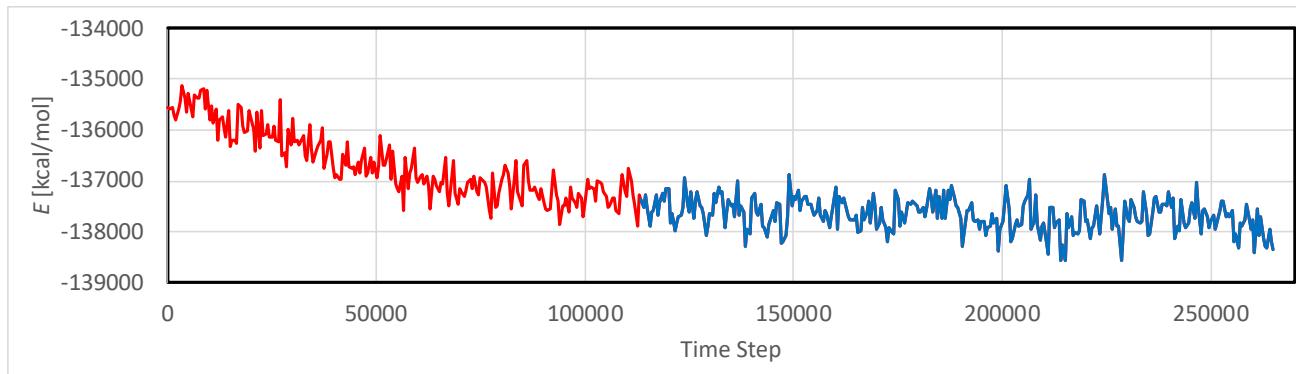


Figure 6.1.1. Plot of time series in `test01.dat`. The section in blue (start time step 115,000) is determined to be in equilibrium for the given input parameters here.

The detailed results are as follows:

```
The start time is 230 (line number) or 115000 (actual time step)
The sample average is -137647.017745 with an error of 72.7370101040329
The sample variance is 26319.1777439837
The total number of segments is 28
The length of each segment is 10 data points
```

THE TEST STATISTICS BELOW

The alpha level you chose is 0.05
 The Z Score used is 1.959964
 The T Score used is 2.069

Mann-Kendall Test for Averages
 The Mann-Kendall Statistic for Averages is -1.89662587126464
 $-1.89662587126464 > -1.959964$, so the test passes.

Mann-Kendall Test for Variance
 The Mann-Kendall Statistic for Variance is 0.632208623754881
 $0.632208623754881 < 1.959964$, so the test passes.

Shapiro-Wilk Test for Normality
 The Test Statistic for Shapiro Wilk is 0.112485291386203
 $0.112485291386203 < 0.916$, so the test passes.

The Von Neumann Test for Trend in Segment Averages
 The Von Neumann Test Statistic is -1.61624112467866
 $-1.61624112467866 > -1.959964$, so the test passes.

The output was saved to a file named “test01_EquCheck.out” under the test0 directory, which can be used by the user to compare.

6.2. Test 2: Unequ.dat

The time series in this test was artificially created (see file test02.dat), which is plotted in Figure 6.2.1. Visual inspection of the figure suggests that the series is not equilibrated.

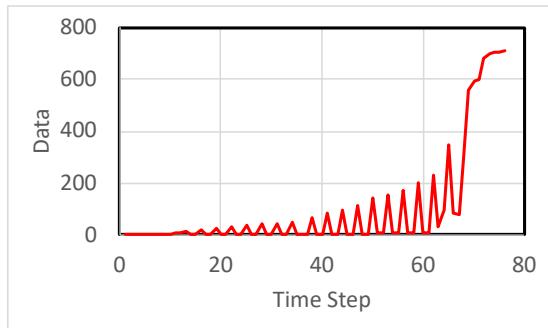


Figure 6.2.1. Plot of the time series in test02.dat, which was found not equilibrated.

This time series failed the tests with the following input parameters without a cap on sample size:

Sample Size = 24
 Time Interval = 1 time units
 Segment Length = 2 time units
 Alpha Level = 0.05
 The User Estimated Start Time = 0 time units

The detailed results are as follows:

Error. It seems equilibration may not have occurred. RELEVANT DATA BELOW.

The alpha level you chose is 0.05
 The Z Score used is 1.959964
 The T Score used is 2.069

The Mann-Kendall Test for Segment Means has failed. Refer to the User Guide for more information.

The Mann-Kendall Statistic for Averages is 1.98078422887447
 $1.98078422887447 > 1.959964$, so the test fails.

The Mann-Kendall Test for Segment Variances has failed. Refer to the User Guide for more information.

The Mann-Kendall Statistic for Variance is 2.08642605441444
 $2.08642605441444 > 1.959964$, so the test fails.

The Shapiro-Wilk Test did not occur, as a previous test may have failed.

The Von-Neumann Test did not occur, as a previous test may have failed.

The output was saved to a file named “test02_EquCheck.out” under the test0 directory, which can be used by the user to compare.

6.3. Test 3: Rand().dat

The time series in this test was artificially created using the Rand() function in MS Excel (see file test03.dat), which is plotted in Figure 6.3.1. Visual inspection of the figure suggests that the series is equilibrated.

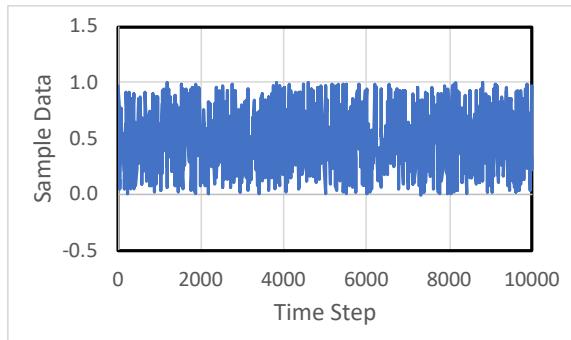


Figure 6.3.1. Plot of the time series in test03.dat, which was found equilibrated starting at the 1st time step.

This time series passed the tests with the following input parameters without a cap on sample size:

Original Sample Size = 24

```
Time Interval = 10 time units
Segment Length = 200 time units
Alpha Level = 0.05
The User Estimated Start Time = 0 time units
```

The detailed results are as follows:

Your data is equilibrated! RELEVANT STATISTICAL INFORMATION BELOW

```
The start time is 1 (line number) or 10 (actual time step)
The sample average is 0.499365102040816 with an error of 0.0251877216933415
The sample variance is 0.00444573369634354
The total number of segments is 49
The length of each segment is 20 data points
```

THE TEST STATISTICS BELOW

```
The alpha level you chose is 0.05
The Z Score used is 1.959964
The T Score used is 2.069
```

```
Mann-Kendall Test for Averages
The Mann-Kendall Statistic for Averages is -0.465471120881711
-0.465471120881711 > -1.959964, so the test passes.
```

```
Mann-Kendall Test for Variance
The Mann-Kendall Statistic for Variance is 1.17229763777616
1.17229763777616 < 1.959964, so the test passes.
```

```
Shapiro-Wilk Test for Normality
The Test Statistic for Shapiro Wilk is 0.000355988165419811
0.000355988165419811 < 0.916, so the test passes.
```

```
The Von Neumann Test for Trend in Segment Averages
The Von Neumann Test Statistic is -0.835344291253944
-0.835344291253944 > -1.959964, so the test passes.
```

The output was saved to a file named “test03_EquCheck.out” under the test0 directory, which can be used by the user to compare.

6.4. Test 4: Rand() + at

The time series in this test was artificially created using the Rand() function in MS Excel (see file test04.dat): $y = \text{Rand}() + 0.0001t$, where t goes from 10 to 10000 with a step size of 10. The time series is plotted in Figure 6.4.1. Visual inspection of the figure suggests that the series is not equilibrated.

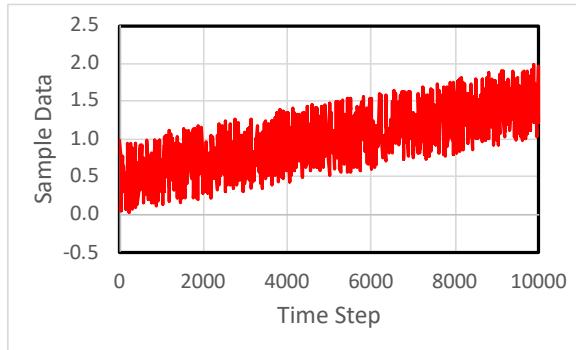


Figure 6.4.1. Plot of the time series in test04.dat, which was found not equilibrated.

This time series failed the tests with the following input parameters without a cap on sample size:

```
Original Sample Size = 24
Time Interval = 10 time units
Segment Length = 200 time units
Alpha Level = 0.05
The User Estimated Start Time = 0 time units
```

The detailed results are as follows:

Error. It seems equilibration may not have occurred. RELEVANT DATA BELOW.

The alpha level you chose is 0.05
 The Z Score used is 1.959964
 The T Score used is 2.069

The Mann-Kendall Test for Segment Means has failed. Refer to the User Guide for more information.

The Mann-Kendall Statistic for Averages is 5.25568082061359
 $5.25568082061359 > 1.959964$, so the test fails.

The Shapiro-Wilk Test did not occur, as a previous test may have failed.

The Von-Neumann Test did not occur, as a previous test may have failed.

The output was saved to a file named “test04_EquCheck.out” under the test0 directory, which can be used by the user to compare.

6.5. Test 5: Rand() + sin(t)

The time series in this test was artificially created using the Rand() function in MS Excel (see file test05.dat): $y = \text{Rand}() + \sin\left(\frac{\pi}{180}\frac{t}{100}\right)$, where t goes from 10 to 10000 with a step size of 10. The time series is plotted in Figure 6.5.1A. Visual inspection of the figure suggests that the series is not equilibrated maybe until near the end of the series.

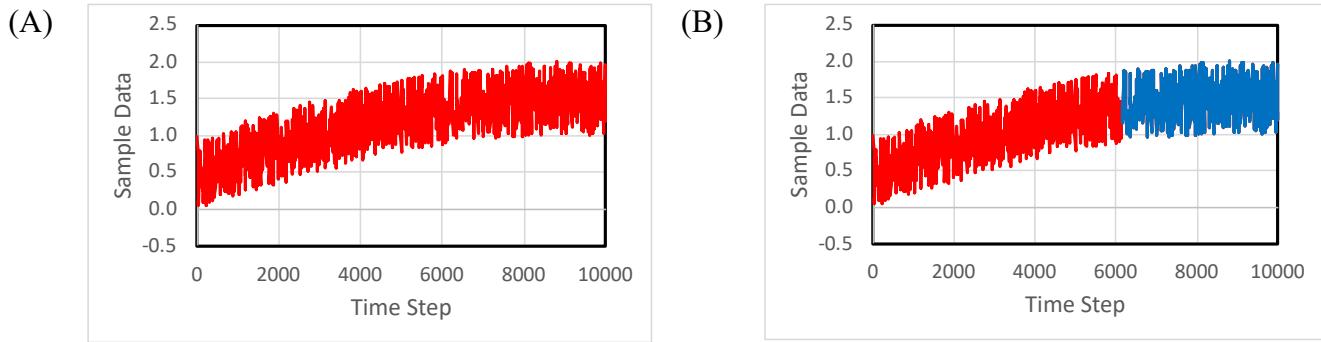


Figure 6.5.1. (A) Plot of the time series in `test05.dat`, which was found not equilibrated using a segment length of 20 data points. (B) The same data set as in (A) was found equilibrated (start time step 6160) using a segment length of 10 data points.

This time series failed the tests with the following input parameters, in particular, segment length of 20 data points, without a cap on sample size (see Fig. 6.5.1A):

```
Original Sample Size = 24
Time Interval = 10 time units
Segment Length = 200 time units
Alpha Level = 0.05
The User Estimated Start Time = 0 time units
```

The detailed results are as follows:

```
Error. It seems equilibration may not have occurred. RELEVANT DATA BELOW.
```

```
The alpha level you chose is 0.05
The Z Score used is 1.959964
The T Score used is 2.069
```

```
The Mann-Kendall Test for Segment Means has failed. Refer to the User Guide for more information.
```

```
The Mann-Kendall Statistic for Averages is 6.04799451216338
6.04799451216338 > 1.959964, so the test fails.
```

```
The Shapiro-Wilk Test did not occur, as a previous test may have failed.
```

```
The Von-Neumann Test did not occur, as a previous test may have failed.
```

However, this time series passed the tests with segment length of 10 data points and no changes to all other parameters (see Fig. 6.5.1B):

```
Original Sample Size = 24
Time Interval = 10 time units
Segment Length = 100 time units
Alpha Level = 0.05
The User Estimated Start Time = 0 time units
```

The detailed results are as follows:

Your data is equilibrated! RELEVANT STATISTICAL INFORMATION BELOW

The start time is 616 (line number) or 6160 (actual time step)
The sample average is 1.44838583333333 with an error of 0.0416503737486228
The sample variance is 0.0113309339392857
The total number of segments is 36
The length of each segment is 10 data points

THE TEST STATISTICS BELOW

The alpha level you chose is 0.05
The Z Score used is 1.959964
The T Score used is 2.069

Mann-Kendall Test for Averages
The Mann-Kendall Statistic for Averages is 1.66174908411375
1.66174908411375 < 1.959964, so the test passes.

Mann-Kendall Test for Variance
The Mann-Kendall Statistic for Variance is 1.93416696675534
1.93416696675534 < 1.959964, so the test passes.

Shapiro-Wilk Test for Normality
The Test Statistic for Shapiro Wilk is 0.0684980129134673
0.0684980129134673 < 0.916, so the test passes.

The Von Neumann Test for Trend in Segment Averages
The Von Neumann Test Statistic is -1.93230240054322
-1.93230240054322 > -1.959964, so the test passes.

The different outcomes demonstrate the importance of segment length (relaxation time) in the tests. The failed and passed outputs were saved, respectively, to files named “test05_EquCheck.out” and “test05_EquCheck_2.out” under the test0 directory, which can be used by the user to compare.

6.6. Test 5: Rand() + ln(t)

The time series in this test was artificially created using the Rand() function in MS Excel (see file test05.dat): $y = \text{Rand}() + 0.1\ln(t)$, where t goes from 10 to 10000 with a step size of 10. The time series is plotted in Figure 6.6.1A. Visual inspection of the figure suggests that the series may be equilibrated after some time.

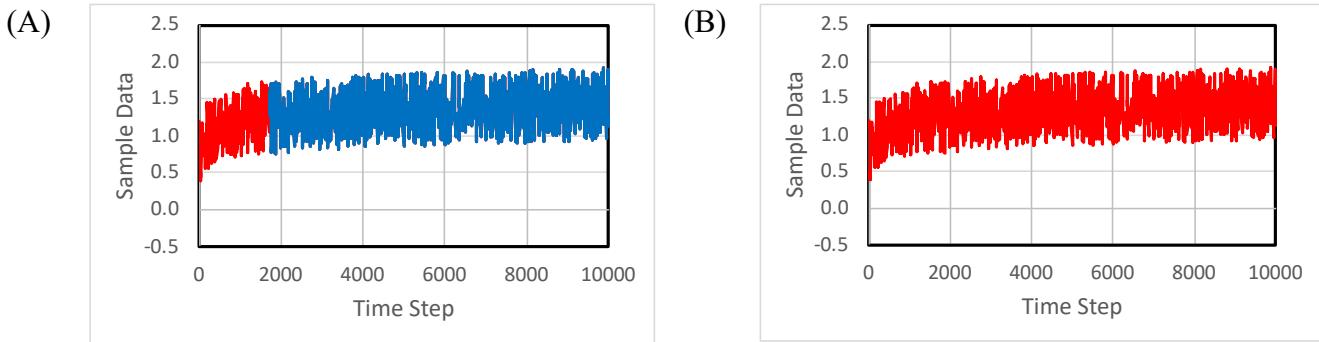


Figure 6.6.1. **(A)** Plot of the time series in `test06.dat`, which was found equilibrated (start time step 1740, indicated in blue) using a segment length of 20 data points. **(B)** The same data set as in (A) was found not equilibrated using a segment length of 40 data points.

This time series passed the tests with the following input parameters, in particular, segment length of 20 data points, without a cap on sample size (see Fig. 6.6.1A):

```
Original Sample Size = 24
Time Interval = 10 time units
Segment Length = 200 time units
Alpha Level = 0.05
The User Estimated Start Time = 0 time units
```

The detailed results are as follows:

Your data is equilibrated! RELEVANT STATISTICAL INFORMATION BELOW

```
The start time is 174 (line number) or 1740 (actual time step)
The sample average is 1.355846375 with an error of 0.0327180868778777
The sample variance is 0.00695620482690706
The total number of segments is 40
The length of each segment is 20 data points
```

THE TEST STATISTICS BELOW

```
The alpha level you chose is 0.05
The Z Score used is 1.959964
The T Score used is 2.069
```

```
Mann-Kendall Test for Averages
The Mann-Kendall Statistic for Averages is 1.86416552971348
1.86416552971348 < 1.959964, so the test passes.
```

```
Mann-Kendall Test for Variance
The Mann-Kendall Statistic for Variance is 0.745666211885391
0.745666211885391 < 1.959964, so the test passes.
```

```
Shapiro-Wilk Test for Normality
The Test Statistic for Shapiro Wilk is 0.112206548092177
0.112206548092177 < 0.916, so the test passes.
```

The Von Neumann Test for Trend in Segment Averages

```
The Von Neumann Test Statistic is -0.903976638732919
-0.903976638732919 > -1.959964, so the test passes.
```

However, this time series failed the tests with segment length of 40 data points and no changes to all other parameters (see Fig. 6.6.1B):

```
Original Sample Size = 24
Time Interval = 10 time units
Segment Length = 400 time units
Alpha Level = 0.05
The User Estimated Start Time = 0 time units
```

The detailed results are as follows:

```
Error. It seems equilibration may not have occurred. RELEVANT DATA BELOW.
```

```
The alpha level you chose is 0.05
The Z Score used is 1.959964
The T Score used is 2.069
```

```
The Mann-Kendall Test for Segment Means has failed. Refer to the User Guide for more
information.
```

```
The Mann-Kendall Statistic for Averages is 3.67105343751401
3.67105343751401 > 1.959964, so the test fails.
```

```
The Shapiro-Wilk Test did not occur, as a previous test may have failed.
```

```
The Von-Neumann Test did not occur, as a previous test may have failed.
```

The different outcomes demonstrate the importance of segment length (relaxation time) in the tests. The failed and passed outputs were saved, respectively, to files named “test06_EquCheck.out” and “test06_EquCheck_2.out” under the testo directory, which can be used by the user to compare.

7. Bibliography

Brown, S. *Measures of Shape: Skewness and Kurtosis*, <https://brownmath.com/stat/shape.htm> (accessed Aug 1, 2018).

Schiferl, S.; Wallace, D. "Statistical errors in molecular dynamics average," *The Journal of Chemical Physics* **1985**, 83, 5203.

Shapiro, S. S.; Wilk, M. B. "An analysis of variance test for normality (complete samples)," *Biometrika* **1965**, 52 (3/4), 591–611.

8. Acknowledgements

This work is supported by the NSF (CHE-1564349), Camille & Henry Dreyfus Foundation (TH-14-028), Research Corporation for Science Advancement (Cottrell SEED Award 25793), and NVIDIA Corporation. This work used XSEDE under grant CHE-140070, supported by NSF grant number ACI-1053575, and NERSC under grant m2495. We thank other Lin Group members, especially Adam Duster, Nara Chon, and Mikias Negussie, at the University of Colorado Denver for providing testing data and supporting us in the development of the program.

9. Appendices

We have implemented a number of tables for statistical analyses in this program:

- The table for converting one-sided percentiles to z-scores was obtained from
<http://users.stat.ufl.edu/~athienit/Tables/Ztable.pdf>
- The table for converting one-sided percentiles to t-scores was obtained from
<http://www.sjsu.edu/faculty/gerstman/StatPrimer/t-table.pdf>
- The table with all the Shapiro-Wilk test values (a -values, n , etc.) was obtained from the original reference (Shapiro and Wilk 1965) – this is the paper on which the statistical test in the code is designed.
- The table with the W -values (the test statistic for the Shapiro-Wilk test) is also found on the original reference (Shapiro and Wilk 1965).

For user's convenience, the above tables are provided below in Tables A1–A4. All copyrights belong to the publishers where the original references are produced.

Table A1. Standard normal probabilities for converting one-sided percentiles to z-scores, for both negative and positive z-scores (Standard Normal Probabilities).

Standard Normal Probabilities										
<i>z</i>	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

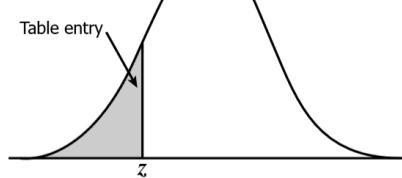


Table entry for z is the area under the standard normal curve to the left of z .

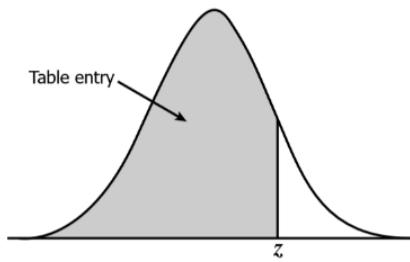


Table entry for z is the area under the standard normal curve to the left of z .

Table A2. Values for the conversion of confidence interval to *t*-score, given degrees of freedom (*t*-table).

t Table

Table A3. Values for a coefficient for n up to 30 for the Shapiro-Wilk test.

$i \backslash n$	2	3	4	5	6	7	8	9	10	
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739	
2	—	·0000	·1677	·2413	·2806	·3031	·3164	·3244	·3291	
3	—	—	—	·0000	·0875	·1401	·1743	·1976	·2141	
4	—	—	—	—	—	·0000	·0561	·0947	·1224	
5	—	—	—	—	—	—	—	·0000	·0399	
$i \backslash n$	11	12	13	14	15	16	17	18	19	20
1	0.5601	0.5475	0.5359	0.5251	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734
2	·3315	·3325	·3325	·3318	·3306	·3290	·3273	·3253	·3232	·3211
3	·2260	·2347	·2412	·2460	·2495	·2521	·2540	·2553	·2561	·2565
4	·1429	·1586	·1707	·1802	·1878	·1939	·1988	·2027	·2059	·2085
5	·0695	·0922	·1099	·1240	·1353	·1447	·1524	·1587	·1641	·1686
6	0.0000	0.0303	0.0539	0.0727	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334
7	—	—	·0000	·0240	·0433	·0593	·0725	·0837	·0932	·1013
8	—	—	—	—	·0000	·0196	·0359	·0496	·0612	·0711
9	—	—	—	—	—	—	·0000	·0163	·0303	·0422
10	—	—	—	—	—	—	—	—	·0000	·0140
$i \backslash n$	21	22	23	24	25	26	27	28	29	30
1	0.4643	0.4590	0.4542	0.4493	0.4450	0.4407	0.4366	0.4328	0.4291	0.4254
2	·3185	·3156	·3126	·3098	·3069	·3043	·3018	·2992	·2968	·2944
3	·2578	·2571	·2563	·2554	·2543	·2533	·2522	·2510	·2499	·2487
4	·2119	·2131	·2139	·2145	·2148	·2151	·2152	·2151	·2150	·2148
5	·1736	·1764	·1787	·1807	·1822	·1836	·1848	·1857	·1864	·1870
6	0.1399	0.1443	0.1480	0.1512	0.1539	0.1563	0.1584	0.1601	0.1616	0.1630
7	·1092	·1150	·1201	·1245	·1283	·1316	·1346	·1372	·1395	·1415
8	·0804	·0878	·0941	·0997	·1046	·1089	·1128	·1162	·1192	·1219
9	·0530	·0618	·0696	·0764	·0823	·0876	·0923	·0965	·1002	·1036
10	·0263	·0368	·0459	·0539	·0610	·0672	·0728	·0778	·0822	·0862
11	0.0000	0.0122	0.0228	0.0321	0.0403	0.0476	0.0540	0.0598	0.0650	0.0697
12	—	—	·0000	·0107	·0200	·0284	·0358	·0424	·0483	·0537
13	—	—	—	—	·0000	·0094	·0178	·0253	·0320	·0381
14	—	—	—	—	—	—	·0000	·0084	·0159	·0227
15	—	—	—	—	—	—	—	—	·0000	·0076

$i \diagdown n$	31	32	33	34	35	36	37	38	39	40
1	0.4220	0.4188	0.4156	0.4127	0.4096	0.4068	0.4040	0.4015	0.3989	0.3964
2	.2921	.2898	.2876	.2854	.2834	.2813	.2794	.2774	.2755	.2737
3	.2475	.2463	.2451	.2439	.2427	.2415	.2403	.2391	.2380	.2368
4	.2145	.2141	.2137	.2132	.2127	.2121	.2116	.2110	.2104	.2098
5	.1874	.1878	.1880	.1882	.1883	.1883	.1881	.1880	.1878	
6	0.1641	0.1651	0.1660	0.1667	0.1673	0.1678	0.1683	0.1686	0.1689	0.1691
7	.1433	.1449	.1463	.1475	.1487	.1496	.1505	.1513	.1520	.1526
8	.1243	.1265	.1284	.1301	.1317	.1331	.1344	.1356	.1366	.1376
9	.1066	.1093	.1118	.1140	.1160	.1179	.1196	.1211	.1225	.1237
10	.0899	.0931	.0961	.0988	.1013	.1036	.1056	.1075	.1092	.1108
11	0.0739	0.0777	0.0812	0.0844	0.0873	0.0900	0.0924	0.0947	0.0967	0.0986
12	.0585	.0629	.0669	.0706	.0739	.0770	.0798	.0824	.0848	.0870
13	.0435	.0485	.0530	.0572	.0610	.0645	.0677	.0706	.0733	.0759
14	.0289	.0344	.0395	.0441	.0484	.0523	.0559	.0592	.0622	.0651
15	.0144	.0206	.0262	.0314	.0361	.0404	.0444	.0481	.0515	.0546
16	0.0000	0.0068	0.0131	0.0187	0.0239	0.0287	0.0331	0.0372	0.0409	0.0444
17	—	—	.0000	.0062	.0119	.0172	.0220	.0264	.0305	.0343
18	—	—	—	—	.0000	.0057	.0110	.0158	.0203	.0244
19	—	—	—	—	—	—	.0000	.0053	.0101	.0146
20	—	—	—	—	—	—	—	—	.0000	.0049
$i \diagdown n$	41	42	43	44	45	46	47	48	49	50
1	0.3940	0.3917	0.3894	0.3872	0.3850	0.3830	0.3808	0.3789	0.3770	0.3751
2	.2719	.2701	.2684	.2667	.2651	.2635	.2620	.2604	.2589	.2574
3	.2357	.2345	.2334	.2323	.2313	.2302	.2291	.2281	.2271	.2260
4	.2091	.2085	.2078	.2072	.2065	.2058	.2052	.2045	.2038	.2032
5	.1876	.1874	.1871	.1868	.1865	.1862	.1859	.1855	.1851	.1847
6	0.1693	0.1694	0.1695	0.1695	0.1695	0.1695	0.1695	0.1693	0.1692	0.1691
7	.1531	.1535	.1539	.1542	.1545	.1548	.1550	.1551	.1553	.1554
8	.1384	.1392	.1398	.1405	.1410	.1415	.1420	.1423	.1427	.1430
9	.1249	.1259	.1269	.1278	.1286	.1293	.1300	.1306	.1312	.1317
10	.1123	.1136	.1149	.1160	.1170	.1180	.1189	.1197	.1205	.1212
11	0.1004	0.1020	0.1035	0.1049	0.1062	0.1073	0.1085	0.1095	0.1105	0.1113
12	.0891	.0909	.0927	.0943	.0959	.0972	.0986	.0998	.1010	.1020
13	.0782	.0804	.0824	.0842	.0860	.0876	.0892	.0906	.0919	.0932
14	.0677	.0701	.0724	.0745	.0765	.0783	.0801	.0817	.0832	.0846
15	.0575	.0602	.0628	.0651	.0673	.0694	.0713	.0731	.0748	.0764
16	0.0476	0.0506	0.0534	0.0560	0.0584	0.0607	0.0628	0.0648	0.0667	0.0685
17	.0379	.0411	.0442	.0471	.0497	.0522	.0546	.0568	.0588	.0608
18	.0283	.0318	.0352	.0383	.0412	.0439	.0465	.0489	.0511	.0532
19	.0188	.0227	.0263	.0296	.0328	.0357	.0385	.0411	.0436	.0459
20	.0094	.0136	.0175	.0211	.0245	.0277	.0307	.0335	.0361	.0386
21	0.0000	0.0045	0.0087	0.0126	0.0163	0.0197	0.0229	0.0259	0.0288	0.0314
22	—	—	.0000	.0042	.0081	.0118	.0153	.0185	.0215	.0244
23	—	—	—	—	.0000	.0039	.0076	.0111	.0143	.0174
24	—	—	—	—	—	—	.0000	.0037	.0071	.0104
25	—	—	—	—	—	—	—	—	.0000	.0035

Table A4. Error (α -level) values for the Shapiro-Wilk test, given the sample size n up to 50.

n	Level								
	0·01	0·02	0·05	0·10	0·50	0·90	0·95	0·98	0·99
3	0·753	0·756	0·767	0·789	0·959	0·998	0·999	1·000	1·000
4	·687	·707	·748	·792	·935	·987	·992	·996	·997
5	·686	·715	·762	·806	·927	·979	·986	·991	·993
6	0·713	0·743	0·788	0·826	0·927	0·974	0·981	0·986	0·989
7	·730	·760	·803	·838	·928	·972	·979	·985	·988
8	·749	·778	·818	·851	·932	·972	·978	·984	·987
9	·764	·791	·829	·859	·935	·972	·978	·984	·986
10	·781	·806	·842	·869	·938	·972	·978	·983	·986
11	0·792	0·817	0·850	0·876	0·940	0·973	0·979	0·984	0·986
12	·805	·828	·859	·883	·943	·973	·979	·984	·986
13	·814	·837	·866	·889	·945	·974	·979	·984	·986
14	·825	·846	·874	·895	·947	·975	·980	·984	·986
15	·835	·855	·881	·901	·950	·975	·980	·984	·987
16	0·844	0·863	0·887	0·906	0·952	0·976	0·981	0·985	0·987
17	·851	·869	·892	·910	·954	·977	·981	·985	·987
18	·858	·874	·897	·914	·956	·978	·982	·986	·988
19	·863	·879	·901	·917	·957	·978	·982	·986	·988
20	·868	·884	·905	·920	·959	·979	·983	·986	·988
21	0·873	0·888	0·908	0·923	0·960	0·980	0·983	0·987	0·989
22	·878	·892	·911	·926	·961	·980	·984	·987	·989
23	·881	·895	·914	·928	·962	·981	·984	·987	·989
24	·884	·898	·916	·930	·963	·981	·984	·987	·989
25	·888	·901	·918	·931	·964	·981	·985	·988	·989
26	0·891	0·904	0·920	0·933	0·965	0·982	0·985	0·988	0·989
27	·894	·906	·923	·935	·965	·982	·985	·988	·990
28	·896	·908	·924	·936	·966	·982	·985	·988	·990
29	·898	·910	·926	·937	·966	·982	·985	·988	·990
30	·900	·912	·927	·939	·967	·983	·985	·988	·990
31	0·902	0·914	0·929	0·940	0·967	0·983	0·986	0·988	0·990
32	·904	·915	·930	·941	·968	·983	·986	·988	·990
33	·906	·917	·931	·942	·968	·983	·986	·989	·990
34	·908	·919	·933	·943	·969	·983	·986	·989	·990
35	·910	·920	·934	·944	·969	·984	·986	·989	·990
36	0·912	0·922	0·935	0·945	0·970	0·984	0·986	0·989	0·990
37	·914	·924	·936	·946	·970	·984	·987	·989	·990
38	·916	·925	·938	·947	·971	·984	·987	·989	·990
39	·917	·927	·939	·948	·971	·984	·987	·989	·991
40	·919	·928	·940	·949	·972	·985	·987	·989	·991
41	0·920	0·929	0·941	0·950	0·972	0·985	0·987	0·989	0·991
42	·922	·930	·942	·951	·972	·985	·987	·989	·991
43	·923	·932	·943	·951	·973	·985	·987	·990	·991
44	·924	·933	·944	·952	·973	·985	·987	·990	·991
45	·926	·934	·945	·953	·973	·985	·988	·990	·991
46	0·927	0·935	0·945	0·953	0·974	0·985	0·988	0·990	0·991
47	·928	·936	·946	·954	·974	·985	·988	·990	·991
48	·929	·937	·947	·954	·974	·985	·988	·990	·991
49	·929	·937	·947	·955	·974	·985	·988	·990	·991
50	·930	·938	·947	·955	·974	·985	·988	·990	·991