

## Steady- and Transient-State Detection

R. Russell Rhinehart

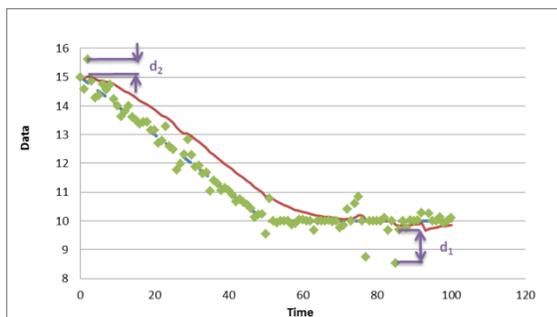
### Introduction

This description is extracted from Nonlinear Regression Modeling for Engineering Applications: Modeling, Model Validation, and Enabling Design of Experiments, by R. R. Rhinehart, John Wiley & Sons, on target for release in September 2016.

The method of Cao and Rhinehart uses a ratio of two variances, as measured on the same set of data by two methods. Figure 1 illustrates the concept (Huang, 2013). The dotted line represents the true trend of a process. The value starts at 15, ramps to a value of 10 at a time of 50, and then holds steady. The diamond markers about that trend represent the measured data. The dashed line, the true trend, is unknowable, only the measurements can be known and they are infected with noise-like fluctuations, masking the truth.

The method first calculates a filtered value of the process measurements, indicated by the curved line that lags behind the data. Then the variance in the data is measured by two methods. The deviation indicated by  $d_2$  in the upper left of the figure is the difference between measurement and the filtered trend. The deviation indicated by  $d_1$  in the lower right is the difference between sequential data measurements.

If the process is at SS, as illustrated in the 80 to 100 time period,  $X_f$  is almost the middle of the data. Then a process variance,  $\sigma^2$ , estimated by  $d_2$  will ideally be equal to  $\sigma^2$  estimated by  $d_1$ . Then the ratio of the variances,  $R = \frac{\sigma^2_{d_2}}{\sigma^2_{d_1}}$  will be approximately equal to unity,  $R = \frac{\sigma^2_{d_2}}{\sigma^2_{d_1}} \cong 1$ . Alternately, if the process is in a TS, then  $X_f$  is not the middle of data, the filtered value lags behind, and the variance as measured by  $d_2$  will be much larger than the variance as estimated by  $d_1$ ,  $\sigma^2_{d_2} \gg \sigma^2_{d_1}$ , and ratio will be much greater than unity,  $R = \frac{\sigma^2_{d_2}}{\sigma^2_{d_1}} \gg 1$ .



**Figure 1: Illustration actual process (dashed line), noisy measurements (diamond markers), filtered data (solid line), and deviations**

## Equations

To minimize computational burden, in this method a filtered value (not an average) provides an estimate of the data mean:

$$X_{f,i} = \lambda_1 X_i + (1 - \lambda_1) X_{f,i-1} \quad (1)$$

$X$  = the process variable

$X_f$  = Filtered value of  $X$

$\lambda_1$  = Filter factor

$i$  = Time sampling index

The first method to obtain a measure of the variance uses an exponentially weighted moving “variance” (another first-order filter) based on the difference between the data and the “average”:

$$v^2_{f,i} = \lambda_2 (X_i - X_{f,i-1})^2 + (1 - \lambda_2) v^2_{f,i-1} \quad (2)$$

$v^2_{f,i}$  = Filtered value of a measure of variance based on differences between data and filtered values

$v^2_{f,i-1}$  = Previous filtered value

Equation (2) is a measure of the variance to be used in the numerator or the ratio statistic. The previous value of the filtered measurement is used instead of the most recently updated value to prevent autocorrelation from biasing the variance estimate,  $v^2_{f,i}$ , keeping the equation for the ratio simple.

The second method to obtain a measure of variance is an exponentially weighted moving “variance” (another filter) based on sequential data differences:

$$\delta^2_{f,i} = \lambda_3 (X_i - X_{i-1})^2 + (1 - \lambda_3) \delta^2_{f,i-1} \quad (3)$$

$\delta^2_{f,i}$  = Filtered value of a measure of variance

$\delta^2_{f,i-1}$  = Previous filtered value

This will be the denominator measure of the variance.

The ratio of variances, the R-statistic, may now be computed by the following simple equation:

$$R = \frac{(2-\lambda_1)v^2_{f,i}}{\delta^2_{f,i}} \quad (4)$$

The calculated value is to be compared to its critical values to determine SS or TS. Neither Equations (2) nor (3) compute the variance. They compute a measure of the variance.

Accordingly, the  $(2 - \lambda_1)$  coefficient in Equation (4) is required to scale the ratio, to represent the variance ratio.

The essential assignment statements for Equations (1) to (4) are:

```
nu2f := L2 * (measurement - xf) ^ 2 + cL2 * nu2f
xf := L1 * measurement + cL1 * xf
delta2f := L3 * (measurement - measurement_old) ^ 2 + cL3 * delta2f
measurement_old := measurement
R_Filter := (2 - L1) * nu2f / delta2f
```

The coefficients L1, L2, and L3 represent the lambda values, and the coefficients cL1, cL2, and cL3 represent the complementary values.

The five computational lines of code of this method require direct, no-logic, low storage, and low computational operation calculations. In total there are four variables and seven coefficients to be stored, ten multiplication or divisions, five additions, and two logical comparisons per observed variable.

Being a ratio of variances the statistic is scaled by the inherent noise level in the data. It is also independent of the dimensions chosen for the variable.

If the process is at steady state then the R-statistic will have a distribution of values near unity. Alternately, if the process is not at steady state then the filtered value will lag behind the data, making the numerator term larger than the denominator, and the ratio will be larger than unity.

### **Coefficient, Threshold, and Sample Frequency Values**

For simplicity and for balancing speed of response with surety of decision, and robustness to noiseless periods, use filter values of  $\lambda_1=\lambda_2=0.1$  and  $\lambda_3=0.05$ . However, other users have recommended alternate values to optimize speed of response and Type-I and type II errors. The method works well for diverse combinations of filter coefficient values within the range of 0.05 to 0.2.

If R-calculated  $> 2.5$ , “reject” steady state with fairly high confidence, accept that the process is in a transient state. Alternately, if R-calculated  $< .9$ , “accept” that the process is at steady-state, reject that it is in a transient state with fairly high confidence. If in-between values for R-calculated, hold the prior SS or TS state, because there is no confidence in changing the declaration.

The filter factors in Equations (1), (2) and (3) can be related to the number of data (the length of the time window) effectively influencing the average or variance calculation. Based on a first-order decay, roughly, the number of data in the window of observation is about  $3.5/\lambda$ .

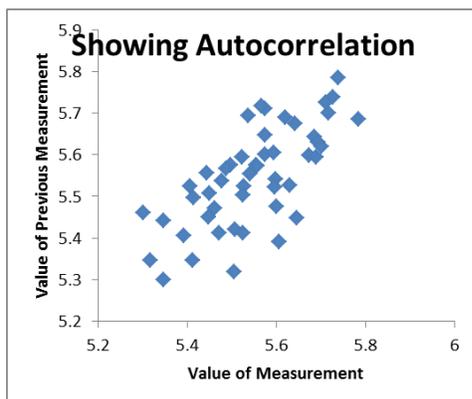
Larger  $\lambda$  values mean that fewer data are involved in the analysis, which has a benefit of reducing the time for the identifier to catch up to a process change, reducing the average run length (ARL). But, larger  $\lambda$  values has an undesired impact of increasing the variability on the statistic, confounding interpretation. The reverse is true: Lower  $\lambda$  values undesirably increase the ARL to detection, but increase precision (minimizing statistical errors).

A faster sampling frequency will speed up the recognition. But, autocorrelation is an issue.

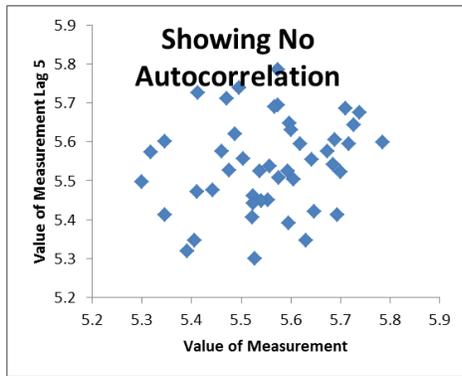
The basis for this method presumes that there is no autocorrelation in the time-series process data when at SS. Autocorrelation means that if a measurement is high (or low) the subsequent measurement will be related to it. For example, if a real process event causes a temperature measurement to be a bit high, and the event has persistence, then the next measurement will also be influenced by the persisting event and will also be a bit high. Autocorrelation could be related to control action, thermal inertia, noise filters in sensors, etc. Autocorrelation would tend to make all R-statistic distributions shift to the right, requiring a reinterpretation of critical values for each process variable.

It is more convenient to choose a sampling interval that eliminates autocorrelation, than to model and compensate for autocorrelation in the test statistic. A plot of the current process measurement versus the previous sampling of the process measurement over a sufficiently long period of time (equaling several time constants) at steady state is required to establish the presence/absence of autocorrelation. To detect autocorrelation, visually choose a segment of data that is at steady state, and plot the PV value vs. its prior value.

Figure 2 plots data with a lag of 1 sample (a measurement vs. the prior measurement) and shows autocorrelation. Figure 3 plots the same data but with a lag of 5 samples, and shows zero autocorrelation.



**Figure 2: Data showing autocorrelation**



**Figure 3: Data showing no autocorrelation when the interval is 5 samplings**

Generally, hydraulics (flow rates, and levels) come to SS faster than gas pressure in large volumes, and faster than temperature or composition. Further, noise in flow rate and level measurements will have little persistence; however, thermal inertia of temperature sensors may extend autocorrelation. So, a sampling interval for one might not be what is needed for another. The user might want to separate the process attributes into a set of hydraulic variables, another set of thermal and composition inventory (T, P, x) variables, and monitor the hydraulic state of the process and the thermal and composition states separately.

Summarizing: Use  $\lambda_1=\lambda_2=0.1$  and  $\lambda_3=0.05$ . Use R-critical = 2.5 to reject SS (accept TS), and R-critical = 0.9 to accept SS (reject TS). Choose a sampling interval to eliminate autocorrelation from a visually-determined SS period.

Figure 4 illustrates the method. The process variable, PV, is connected to the left hand vertical axis (log10-scale) and is graphed with respect to sample interval. Initially it is at a steady state with a value of about 5. At a sample number 200, the PV begins a first-order rise to a value of about 36. At sample number 700, the PV makes a step rise to a value of about 36. The ratio-statistic is attached to the same left-hand axis and shows an initial kick to a high value as variables are initialized, then relaxes to a value that wanders about the unity SS value. When the PV changes at sample 200, the R-statistic value jumps up to values ranging between 4 and 11, which relaxes back to the unity value as the trend hits a steady value at a time of 500. Then when the small PV step occurs at sample 700, the R-value jumps up to about 4, then decays back to its nominal unity range. The SS value is connected to the right-hand vertical axis and has values of wither 0 to 1 that change when the R-value exceeds the two limits, of  $R_{\beta,TS}$  and  $R_{1-\alpha,SS}$ .

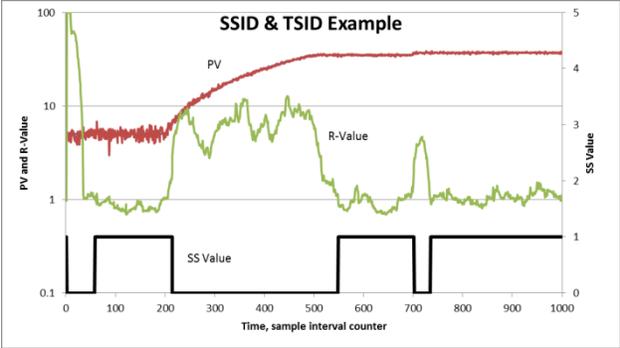


Figure 4: Example