



All Faculty Publications

2013-3

A Steady-State Detection (SSD) Algorithm to Detect Non-Stationary Drifts in Processes

Jeff Kelly

John Hedengren

Brigham Young University, john.hedengren@byu.edu

Follow this and additional works at: <https://scholarsarchive.byu.edu/facpub>

 Part of the [Chemical Engineering Commons](#)

Original Publication Citation

<http://www.sciencedirect.com/science/article/pii/S0959152412002661>

BYU ScholarsArchive Citation

Kelly, Jeff and Hedengren, John, "A Steady-State Detection (SSD) Algorithm to Detect Non-Stationary Drifts in Processes" (2013). *All Faculty Publications*. 1713.

<https://scholarsarchive.byu.edu/facpub/1713>

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

A Steady-State Detection (SSD) Algorithm to Detect Non-Stationary Drifts in Processes

Jeffrey D. Kelly

Industrial Algorithms, 15 St. Andrews Road, Toronto, ON, Canada, M1P 4C3

John D. Hedengren*

Department of Chemical Engineering, Brigham Young University, Provo, UT 84602

Abstract

Detecting windows or intervals of when a continuous process is operating in a state of steadiness is useful especially when steady-state models are being used to optimize the process or plant on-line or in real-time. The term steady-state implies that the process is operating around some stable point or within some stationary region where it must be assumed that the accumulation or rate-of-change of material, energy and momentum is statistically insignificant or negligible. This new approach is to assume the null-hypothesis that the process is stationary about its mean subject to independent and identically distributed random error or shocks (white-noise) with the alternative-hypothesis that it is non-stationary with a detectable and deterministic slope, trend, bias or drift. The drift profile would be typical of a time-varying inventory or holdup of material with imbalanced flows or even an unexpected leak indicating that the process signal is not steady. A probability of being steady or at least stationary over the window is computed by performing a residual Student-t test using the estimated mean of the process signal without any drift and the estimated standard-deviation of the underlying white-noise driving force. There are essentially two settings or options for the method which are the window-length and the Student-t critical value and can be easily tuned for each process signal that are included in the multivariate detection strategy.

Keywords: steady-state, stationarity, random walk with drift, white-noise, hypothesis testing, student-t.

1. Introduction

2 If the process or plant being monitored (passively) and/or optimized (actively) is
3 not at steady-state then applying a steady-state model at that time is obviously not

*Corresponding author

Email addresses: jdkelly@industrialgorithms.ca (Jeffrey D. Kelly),
john.hedengren@byu.edu (John D. Hedengren)

4 suitable given that significant accumulation or rate-of-change of material, energy and
5 momentum violates one of the principle assumptions of the model. Applying the right
6 model at the wrong time will result in more Type I and II errors (false positives, false
7 negatives), biased or inaccurate parameter estimates and ultimately inappropriate de-
8 cisions to be made on how to move the system to be more economical, efficient and
9 effective. Serious violation of the steady-state assumption may result in unstable op-
10 eration when on-line, real-time or closed-loop optimization is applied. Thus, correctly
11 detecting intervals, horizons or windows when steady-state models can most likely be
12 used is an important application. Knowing accurately in time when some processes
13 are steady and others are unsteady can also help to identify and diagnosis potentially
14 coincident abnormal events or symptoms noticed in other areas of the plant such as
15 light-ends flaring, low-pressure steam venting and contaminated liquid effluents.

16 The subject of this work is to highlight a straightforward technique to detect pe-
17 riods of time in the immediate past and present when the continuous process appears
18 to be running in a state of steady-ness or is stationary from which it is reasonable to
19 assume that steady-state models can be implemented for the very near future. Previ-
20 ous work in the area of steady-state detection (SSD) is summarized by Mhamdi et.
21 al. [1] as (a) performing a Student-t test on a linear regressed slope over the time
22 window, (b) performing a Student-t test on two recently computed means with pooled
23 standard-deviations from two adjacent windows and (c) performing an F-test on two
24 recently computed standard-deviations either from two adjacent windows or from the
25 same window but using two different filtered means. Examples of (a) can be found in
26 Holly et. al. [2] and Bethea and Rhinehart [3], examples of (b) in Narasimhan et. al.
27 [4] and Holly et. al. [2] and examples of (c) in Cao and Rhinehart [5] and Mansour
28 and Ellis [6] using pre-specified exponentially-weighted filters with an interesting re-
29 cursive window-based version found in Kim et. al. [7]. The method of Mhamdi et.
30 al. [1] is somewhat similar to (a) except that it uses more sophisticated basis func-
31 tions implemented in wavelet theory instead of employing a slope, trend, bias or drift
32 component used here and as such is more difficult to apply.

33 Most industrial implementations of SSD use a form of (b) known as the mathe-
34 matical theory of evidence (Narasimhan et. al. [4]) usually with another Student-t
35 test on the residuals of the raw signal minus its mean divided by its standard-deviation
36 computed over the number of data values in the window. Unfortunately, the mean and
37 standard-deviation computed by these methods are not corrected for the drift compo-
38 nent as is done in this SSD algorithm below. Hence, the other methods are biased (less
39 accurate) and require more adjustment to minimize Type I and II errors. Although
40 these techniques are easy to understand and implement, it is well-known that they re-
41 quire substantial and subjective tuning or calibration knowing intervals of when the
42 plant is possibly at steady-state (Campos et. al. [8]) and is a perceived drawback.
43 In terms of computational expense, recursive techniques can significantly reduce the
44 computing load but since the eighties with mainframes as the supervisory computers
45 and now with multi-core application servers this is not an issue to consider further es-
46 pecially for the SSD algorithm described here. The SSD algorithm uses insignificant
47 CPU time because it only involves calculation of a mean, standard deviation, and slope.

48 The SSD algorithm presented in this work is also window-based and utilizes the
49 Student-t test to determine if the difference between the process signal value minus its

50 mean is above or below the standard-deviation times its statistical critical value. If less
 51 than, then that time instant or point is steady and if greater than, then it is unsteady
 52 where the aggregation is computed over the window approximating a probability or
 53 frequency of being at steady-state. The details of this algorithm are now presented.

54 2. Steady-State Detection (SDD) Algorithm

55 Our fundamental assumption about the behavior of the underlying system for any
 56 single process signal is to assume that it may be operating with a non-zero slope mul-
 57 tiplied by its relative time within the window defined by the following equation for x_t
 58 as:

$$x_t = mt + \mu + a_t \quad (1)$$

59 where mt is the deterministic drift component, μ is the mean of the hypothetical
 60 stationary process that will also equal the sample mean or arithmetic average over
 61 the time window with zero slope, and a_t is the i.i.d. random error series or white-
 62 noise sequence with zero mean and standard-deviation σ_a . Subscript t is an index
 63 that indicates the cycle at which the sample is collected while mt refers to the slope
 64 multiplied by cycle count t . This is well-known as the “random walk with drift” non-
 65 stationary time-series found in Box and Jenkins [9] which is clearer to see when the
 66 difference x_t is first lagged or time-shifted in the immediate past x_{t-1} as:

$$x_t - x_{t-1} = m + a_t - a_{t-1} \quad (2)$$

67 where $a_t - a_{t-1}$ by definition has an expected value of zero with a standard-deviation
 68 of $2\sigma_a$. This is the simplest type of a non-stationary process and can be used to
 69 model any process with non-constant accumulation or rate-of-change of material, en-
 70 ergy and/or momentum. An example would be a process vessel with a holdup or in-
 71 ventory of material where m would be non-zero with a net flow in or out of the vessel
 72 or even an unexpected leak either at the input, output or inside the vessel itself due
 73 to a loss of integrity in the system. If it is assumed that any process signal (usually a
 74 dependent variable such as a temperature, pressure or concentration) included in the
 75 steady-state detection can accumulate over time, and if this is found to be significant,
 76 then the process variable can be declared to be unsteady or non-stationary. A similar
 77 approach was taken by Kelly [10] to model the non-stationary disturbances caused
 78 by either the input or output flow, depending on if the output or input flows are the
 79 manipulated variables, of a surge vessel. A pure random walk time-series was chosen
 80 to aid in the tuning of PI controllers for improved level control known as level-flow
 81 smoothing. However, in this case the random walk with drift is used to hypothetically
 82 model potentially sustained accumulation or rate-of-change of holdup or level of the
 83 vessel over the window. This can be equally applied to any process variable which can
 84 accumulate other phenomenon such as energy or momentum with respect to time.

¹Unbiased if the noise is i.i.d. else if there are stationary auto-regressive (AR) components (Box and Jenkins [9]) then a biased estimate will result.

85 By first differencing x_t it is possible to unbiasedly¹ estimate the slope m of the drift
 86 component mt as the arithmetic average of $x_t - x_{t-1}$ with n sampled values of x_t in
 87 the window which are equally spaced in time i.e., given a uniform sampling period or
 88 cycle. Slope in the linear regression can also be obtained by minimizing the sum of
 89 squared errors of the difference between the measured and model values, but this was
 90 not applied here. The slope calculation is the arithmetic mean of the first difference in
 91 x_t and is essentially a discrete-time or first-difference calculation of the first-derivative
 92 or rate-of-change of x_t with respect to time $\left(\frac{\partial x}{\partial t}\right)$ and as such m is the direct calculation
 93 or estimate of this accumulation. Obviously if there is detectable accumulation then the
 94 signal is by definition unsteady given that $\left(\frac{\partial x}{\partial t}\right)$ is found to be non-zero. The intercept
 95 μ is obtained by subtracting the term mt from x_t when Equation 1 is rearranged to
 96 $x_t - mt = \mu + a_t$:

$$\mu = \frac{1}{n} \left(\sum_{t=1}^n x_t - m \sum_{i=1}^n t \right) \quad (3)$$

97 Now that there is an estimate of the drift slope m and the mean μ of x_t , the standard-
 98 deviation of the white-noise shocks can be estimated as:

$$\sigma_a = \sqrt{\frac{1}{n-2} \sum_{t=1}^n (x_t - mt - \mu)^2} \quad (4)$$

99 At this point along with a specified Student-t critical or threshold value at a particu-
 100 lar significance level α and degrees-of-freedom n , all of the necessary information is
 101 available to test the null-hypothesis that the process signal is steady or is stationary
 102 about μ :

$$\text{if } |x_t - \mu| \leq t_{crit} \sigma_a \text{ then } y_t = 1 \text{ else } y_t = 0 \quad (5)$$

103 The sum of y_t divided by n is a fraction related to the likelihood that the null hypothesis
 104 is false. This is the fraction of time within the window that the process or plant is
 105 deemed to be at steady-state. A fraction of 0.95 or 95% would indicate that 5% of the
 106 points are not at steady-state for example. A suitable cut-off determination of whether
 107 the process is deemed to be at steady-state depends on the application. A Student-t test
 108 could also be applied on the estimated slope m of the drift component used in previous
 109 work (see point (a) above) but this was not found to be as accurate as computing the
 110 probability over the window i.e., performing n drift-corrected residual Student-t tests
 111 and taking the arithmetic average of y_t .

112 Before proceeding to the results and discussion section, there are two issues that
 113 need to be addressed and they are the window-length or size and how to manage mul-
 114 tivariate processes. The guideline for the window-length is to set it at some number
 115 of samples equivalent to greater than say three (3) to five (5) times the time-constant
 116 of the process variable divided by the sampling time-period. This implies that some
 117 prior knowledge of the process is necessary but it is more qualitative than quantitative.
 118 For example, if the time-constant is circa 30-minutes (the time to reach around 63%
 119 or $(1 - e^{-1})$ of its steady-state gain value for a first-order process) and the sampling
 120 takes place every 2-minutes then the window-size should be between $3 \times \frac{30}{2} = 45$ to

121 $5 \times \frac{30}{2} = 75$ number of samples in the window. This is a typical approach used in indus-
 122 try for two-step²on-line optimization installations. The window-length should not be
 123 too short because the process will not have time to reach some level of stability and the
 124 steady-state probability will always be low precluding the use of steady-state models
 125 to help improve the profitability and performance of the plant. Too long, and multi-
 126 ple intervals of unsteady-state behavior within the longer window may conclude that
 127 the signal is steady when in fact it is not. These are well-known effects referred to as
 128 low/high frequency aliasing and under/over-sampling. Another possible short-coming
 129 of this approach is that false indications of steady-state may occur at the peak or valley
 130 of an oscillating process. With the horizon window centered over the peak, a slope
 131 of $m = 0$ and points within the confidence interval of Equation 5 will result. In these
 132 cases, the SSD will indicate periodic acceptance proportional to the frequency of the
 133 oscillation.

134 To manage multiple process signals where collectively they determine whether a
 135 system is steady, the same approach is used as found in the gross-error detection liter-
 136 ature (Narasimhan and Jordache [11]) which was also equivalently presented in Man-
 137 sour and Ellis [6] to handle multivariate systems. Essentially, the individual signifi-
 138 cance level α' is reduced or corrected from the overall significance level α derived
 139 from the well-known Sidak inequality as:

$$\alpha' = 1 - (1 - \alpha)^{\frac{1}{k}} \quad (6)$$

140 where k is the number of key process variables selected to be included in determining
 141 if the process or plant is steady or exhibits some level of stationarity where a number
 142 between three (3) and thirty (30) is reasonable. This means that α' will be smaller
 143 compared to α if k is greater than 1.0 and will result in a larger Student-t critical value
 144 for the same number of degrees-of-freedom n for an individual signal. The types of
 145 process signals included in the set of key variables should be some mix of manipulated
 146 (independent) and controlled (dependent) variables. Although strictly speaking, the
 147 k variables should be independent from each other for Equation 6 to be valid, it is a
 148 fair correction for these purposes to at least address multivariate systems in some way.
 149 Using several key process signals is also useful to identify individual signals which if
 150 always steady when the rest are unsteady, can be an indication that its window-length
 151 is too long. Or, if it is always unsteady when the rest are steady, may imply that its
 152 window-length is too short.

153 For this study, the algorithm is applied to the data in offline batch segments as
 154 demonstrated in the example applications. In these examples, there is no overlap of the
 155 analyzed time horizons. When the null hypothesis is not rejected, all of the points in the
 156 time window are deemed to be at steady-state. An alternative application for real-time
 157 systems is to process the data online as new measurements arrive. This moving horizon
 158 approach would enable real-time monitoring of the process steadiness. Additionally,
 159 multiple horizons could be processed at each sampling interval to determine the degree
 160 of steadiness over multiple time periods.

²Two-step meaning that there is first an estimation run which performs parameterization and reconcili-
 ation before the optimization run similar to "bias-updating" in on-line model-based control applications.

161 **3. Results and Discussion**

162 The testing of this SSD algorithm entails simulating a process signal with a mean
 163 of zero ($\mu = 0$) and superimposed white-noise (a_t) generated from the code found in
 164 Ahrens et. al. [12] with a standard-deviation specified as 1.0 ($\sigma_a = 1$) for simplicity.
 165 A window-size of 120 samples or time-periods is used where the sampling instant
 166 is assumed to be one-minute in duration simulating two-hours of real-time. Three
 167 disturbance model structures were used to generate the data sequences or time-series
 168 which are a stationary periodic cycle with $m \sin(t) \neq 0$, a non-stationary drift with
 169 $mt \neq 0$ and a stationary stochastic process with $ARMA(p = 1, q = 1)$. For the first two
 170 disturbance models involving m the value value is varied from 0.001, 0.01, 0.1, 1.0
 171 and 10.0 in order to assess the sensitivity to the signal-to-noise ratio which also varies
 172 from 0.001 to 10.0 given the fixed choice of $\sigma_a = 1$. For the auto-regressive part of the
 173 model found below in Equation 7, the ϕ_1 parameter is varied within the set of 0, -0.7,
 174 0.9, -0.95 and -0.99 and for the moving-average part the θ_1 parameter is fixed at either
 175 0 or -0.5.

$$x_t = mt + \mu + \frac{1 + \theta_1 z^{-1}}{1 + \phi_1 z^{-1}} a_t \quad (7)$$

176 where z^{-1} represents the lagging of one sampling instant in the past. As ϕ_1 approaches
 177 1.0 then this becomes the most basic form of a non-stationary process (random walk)
 178 and when both ϕ_1 and θ_1 are zero then it reduces again to Equation 1. In order to
 179 confirm the standard-deviation estimate of the white-noise ($\sigma_a = 1$) with $m = 0$, and
 180 $\theta_1 = 0$, the window-lengths are varied to 120, 1200, 12000 and 24000 yielding 1.208,
 181 1.034, 1.010 and 0.999 respectively. Since these estimates are close to 1.0, this con-
 182 firms that the driving force for the simulation is sufficiently distributed as random error.
 183 It also verifies that the calculation of white-noise standard-deviation found in Equation
 184 4 is acceptable as well. Tables 1 and 2 show the simulated probabilities in parentheses
 185 at two different Student-t critical values of 2.0 and 3.0. The value 2.0 typically repre-
 186 sents a 5% significance level and 3.0 is typical of a 0.5% significance. Table 1 using
 187 $m \sin(t)$ is purposely chosen to be a stationary but cyclic deterministic type of process
 188 to show that the SSD algorithm has no reason to reject the null-hypothesis that the sys-
 189 tem is at steady-state although the signal is oscillating within the window but it is not
 190 static. All of the cells of Table 1 are sufficiently close to 95% and 99% confirming that
 191 the process signal is statistically stationary.

192 Using the same random seed as for Table 1, Table 2 exhibits the same results for the
 193 first row as in Table 1. The non-stationary disturbance is detected to be unsteady when
 194 m is greater than 0.01 with only white-noise (second column) and as colored-noise
 195 ($ARMA(1, 1)$) is added unsteady-state operation is detected for m as low as 0.001 when
 196 the colored-noise also tends to approach non-stationarity (fifth and sixth columns).

197 The sensitivity of the SSD to identify unsteady-state activity when a drift is injected
 198 into the signal response has been demonstrated at least for the white and colored-noise
 199 series considered here. As the drift component magnitude increases it gets easier for
 200 the technique to declare the system unsteady (low probability of being steady) espe-
 201 cially when mt is consistently near or above the standard-deviation of the white-noise
 202 driving force. This is easily seen in Figure 1 where both $m = 1$ (solid line) and $m = 0.1$
 203 (dotted line) are plotted with only white-noise. The larger m exhibits an obvious drift

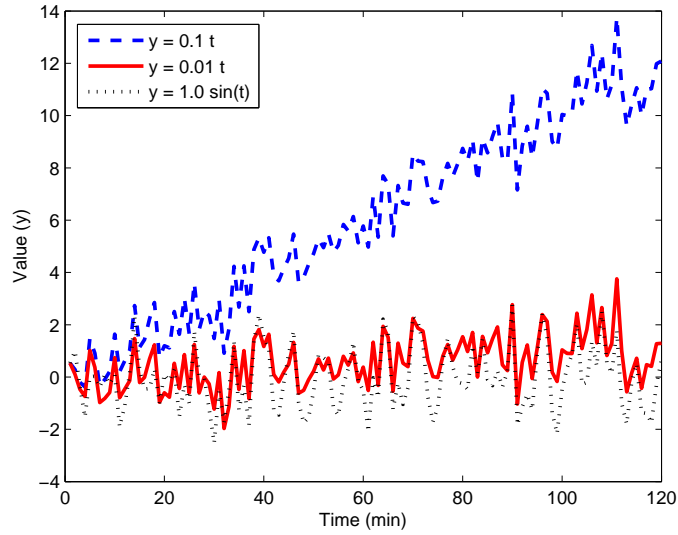


Figure 1: Plot of $1.0\sin(t)$, $0.01t$, and $0.1t$ over the window with $n = 120$.

Table 1: Simulated probability (%) results with $m\sin(t)$ using two Student-t critical values.

m	$\theta_1 = 0.0$ $\phi_1 = 0.0$	$\theta_1 = -0.5$ $\phi_1 = -0.7$	$\theta_1 = -0.5$ $\phi_1 = -0.9$	$\theta_1 = -0.5$ $\phi_1 = -0.95$	$\theta_1 = -0.5$ $\phi_1 = -0.99$
0.0	(95.8,100)	(95.8,100)	(95.8,100)	(95.0,100)	(97.5,100)
0.001	(95.8,100)	(95.8,100)	(95.8,100)	(95.0,100)	(97.5,100)
0.01	(95.8,100)	(95.8,100)	(95.8,100)	(95.0,100)	(97.5,100)
0.1	(96.7,100)	(96.7,100)	(95.8,100)	(95.0,100)	(96.7,100)
1.0	(97.5,100)	(97.5,100)	(95.8,100)	(97.5,100)	(96.7,100)
10.0	(100,100)	(100,100)	(100,100)	(99.2,100)	(99.2,100)

Table 2: Simulated probability (%) results with mt using two Student-t critical values.

m	$\theta_1 = 0.0$ $\phi_1 = 0.0$	$\theta_1 = -0.5$ $\phi_1 = -0.7$	$\theta_1 = -0.5$ $\phi_1 = -0.9$	$\theta_1 = -0.5$ $\phi_1 = -0.95$	$\theta_1 = -0.5$ $\phi_1 = -0.99$
0.0	(95.8,100)	(95.8,100)	(95.8,100)	(95.0,100)	(97.5,100)
0.001	(95.8,100)	(95.8,100.0)	(95.0,99.2)	(87.5,97.5)	(77.5,89.2)
0.01	(92.5,98.3)	(58.3,81.7)	(38.5,52.5)	(31.7,40.8)	(23.3,36.7)
0.1	(20.0,26.7)	(0.1,10.8)	(0.0,10.0)	(0.0,10.8)	(0.0,19.2)
1.0	(0.0,0.0)	(0.0,0.1)	(0.0,0.1)	(0.0,0.1)	(0.0,17.5)
10.0	(0.1,0.1)	(0.0,0.0)	(0.0,0.0)	(0.0,0.1)	(0.0,16.3)

204 up whereas the smaller m has a more subtle trend up and cannot be declared as unsteady
 205 although it does have a lower probability than $m = 0.001$ in column two indicating an
 206 increased residual violation of the stationarity assumption. The term $1.0\sin(t)$ (second
 207 dotted line) is also plotted to not only show it is stationary as confirmed in Table 1
 208 but also to highlight the slight but noticeable drift up of the solid line as the sample
 209 number increases. And expectantly, when the window-length is increased from 120 to
 210 480 samples probabilities of (49.0,70.0) are obtained confirming that the signal is non-
 211 stationary though requiring more time to statistically detect that it is unsteady. This
 212 is not an unusual observation given that it is well-accepted that subtle perturbations
 213 require more sample or data points.

214 3.1. Multivariate Case Study

215 A simple model of a jacketed continuously stirred tank reactor (CSTR) is used to
 216 demonstrate the SSD algorithm. The problem has been used extensively in the litera-
 217 ture to benchmark new techniques because of some unique characteristics that pose a
 218 variety of desirable challenges [13]. One challenge is the nonlinearity of the system
 219 due to the exothermic first-order reaction. The exponential dependency on tempera-
 220 ture causes order of magnitude differences in reaction rates depending on the reactor
 221 temperature. Above a jacket temperature of approximately 305 K, the CSTR enters a
 222 sustained oscillation of temperature run-away followed by reaction quenching and con-
 223 centration build-up of A . Once the concentration C_A reaches a sufficiently high level
 224 the temperature runs away, leading to the next cycle. For this case study, the CSTR is
 225 perturbed by adjusting the jacket temperature T_c but does not become unstable as men-
 226 tioned above. The nonlinear model demonstrates that the SSD algorithm is applicable
 227 to multivariate processes with strong nonlinearities.

228 The CSTR model consists of a feed stream of pure A at concentration $C_{A,i}$ and inlet
 229 temperature T_i . The reactor is well mixed and produces product B with an exothermic
 230 first-order reaction. The reactor temperature and extent of reaction are controlled by
 231 manipulating the cooling jacket temperature T_c with negligible dynamics for the speed
 232 of cooling jacket temperature response. The variables for this CSTR model are shown
 233 in Table 3 and the equations are shown in Table 4.

Table 3: CSTR Parameters and Variables

<i>Manipulated variable</i>			
Symbol	Description	Nominal Value	Units
T_c	Jacket Temperature	300	K
<i>State variables</i>			
Symbol	Description	Nominal Value	Units
C_A	Concentration of A in the reactor	0.877	$\frac{mol}{m^3}$
T	Temperature of the reactor	324.48	K
<i>Other parameters</i>			
Symbol	Description	Value	Units
$C_{A,i}$	Concentration of A in the feed	1.0	$\frac{mol}{m^3}$
C_p	Heat capacity of the liquid	0.239	$\frac{J}{kg K}$
E_a	Activation energy	7.28e4	$\frac{J}{mol}$
ΔH_r	Energy of reaction	5x10 ⁴	$\frac{J}{mol}$
k_0	Pre-exponential factor	7.2e10	$\frac{mol}{m^3 min}$
R	Universal gas constant	8.31451	$\frac{J}{mol K}$
ρ	Mixture density	1000.0	$\frac{kg}{m^3}$
T_i	Feed temperature	350.0	K
q	Feed flow rate	100.0	$\frac{mol}{min}$
V	Volume of the reactor	100.0	m^3

Table 4: CSTR Model Equations

<p>Component balance on A</p> $V \frac{\partial C_A}{\partial t} = qC_{A,i} - qC_A - k_0 C_A V \exp\left(-\frac{E}{RT}\right)$ <p>Energy balance</p> $\rho C_p V \frac{\partial T}{\partial t} = \rho C_p q (T_i - T) + \Delta H_r k_0 C_A V \exp\left(-\frac{E}{RT}\right) + UA (T_c - T)$

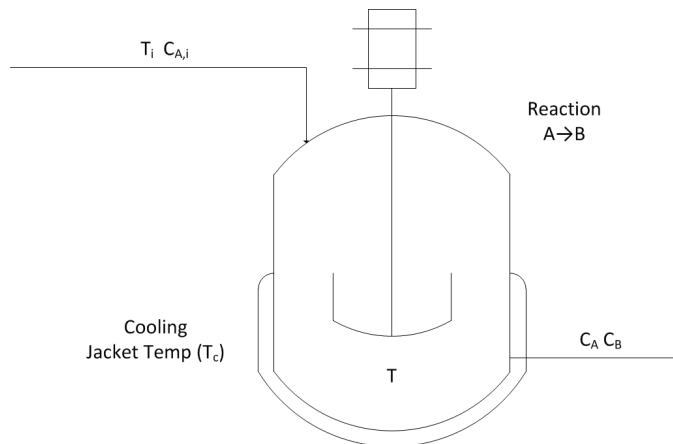


Figure 2: Diagram of the exothermic CSTR with first-order reaction kinetics.

234 The process simulation records data every 1-second over a total time of 40-minutes
 235 for a total of 2400 samples. A step test determined that the time constant is ap-
 236 proximately 1-minute for C_A and 45-seconds for T . According to the guidance pro-
 237 vided earlier, a window of 3 to 5 time constants (using the dominant time constant)
 238 is selected for analyzing the probability that the process is at steady-state. In this
 239 case, a time window of 5 time constants or 5-minutes is selected. Each time win-
 240 dows includes 300 samples for both C_A and T . In the prior examples, the Student-t
 241 critical values of 2.0 and 3.0 were used to determine the steady-state probabilities.
 242 Because this system involves more than one variable, the Sidak inequality suggests
 243 Student-t values of 2.25 and 3.04 for the 5% and 0.5% significance levels, respec-
 244 tively. The cooling temperature is initially lowered from 300K to 290K for a period
 245 of 10-minutes followed by a step back to 300K for another period of 10-minutes. Fol-
 246 lowing these step changes, T_c begins oscillating with a period of 3-seconds for 7-
 247 minutes before returning to the constant value of 300K for the remainder of the total
 248 40-minutes. Random state ($\sigma(\omega_{C_A}) = 0.005, \sigma(\omega_T) = 0.05$) and measurement noise
 249 ($\sigma(v_{C_A}) = 0.02, \sigma(v_T) = 0.2$) are added at each sample point after the equations in
 250 Table 4 are integrated forward in time as shown in Equation 8.

$$x[t + 1] = f(x[t], u[t]) + \omega \quad (8a)$$

$$y[t] = g(x[t], u[t]) + v \quad (8b)$$

252 with $x[t]$ and $y[t]$ being the state and measurement vectors, respectively, for both C_A and
 253 T . The vector $u[t]$ includes all exogenous inputs and t is the cycle index. Functions f
 254 and g are the nonlinear state and measurement functions with g simplifying to $(C_A \ T)^T$
 255 for this example problem.

256 Windows 2, 7, and 8 have the highest probability ($> 90\%$) of being at steady-state
 257 for both C_A and T above the minimum probability limit. Windows 4, 5, and 6 have
 258 either C_A or T greater than a 90% probability to a 5% significance level (first number

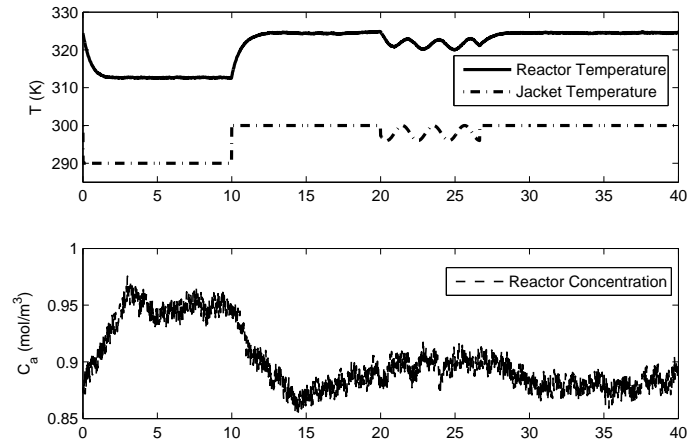


Figure 3: Simulated operational data for the CSTR. Periods of unsteady behavior in reactor concentration and temperature are observed due to steps and sinusoidal fluctuations in the jacket cooling temperature.

Table 5: Simulated probability (%) results with two Student-t critical values. The first probability represents a 5% significance level while the second represents a 0.5% significance.

Index	Time Period	Probability of C_A at Steady-State	Probability of T at Steady-State
1	0 to 5-minutes	(56.7,80.0)	(17.7,100.0)
2	5 to 10-minutes	(99.3,100.0)	(100.0,100.0)
3	10 to 15-minutes	(28.3,39.7)	(37.3,100.0)
4	15 to 20-minutes	(86.7,96.3)	(100.0,100.0)
5	20 to 25-minutes	(97.7,100.0)	(61.3,100.0)
6	25 to 30-minutes	(54.3,78.3)	(100.0,100.0)
7	30 to 35-minutes	(99.3,100.0)	(100.0,100.0)
8	35 to 40-minutes	(97.7,100.0)	(100.0,100.0)

259 in the parenthesis). The other two time windows (1 and 3 bolded) are deemed to not be
260 at steady-state because both C_A and T fail to meet a minimum probability of steadiness
261 ($\leq 90\%$). These results are also visually consistent with Figure 3 as the two largest
262 step changes occur in time windows 1 and 3.

263 3.2. Scale-up to Large-Scale and Complex Systems

264 One concern with any data analysis technique is the scale-up to large-scale sys-
265 tems. In this regard, the CPU time requirements for this SSD algorithm are negligible
266 because it only involves calculation of a mean, slope, and standard deviation for each
267 measurement. For practical purposes, it may be reasonable to select an appropriate
268 cut-off probability value such as 90% for instance and/or to take an average of the two
269 low and high probability estimates (corresponding to the low and high Student-t critical
270 values) and apply the cut-off to this average. This is left as an implementation issue
271 where it can be used to assist in the tuning or aligning of the quantitative results with
272 the qualitative expectations. Although the details cannot be released, an application us-
273 ing this SSD technique been implemented in a fully integrated oil-refinery where SSD
274 was considered to be a key plant, process or performance indicator (KPI) and was used
275 to help isolate temporal root causes to process incidents.

276 An additional application of this technique is in identification of historical data
277 windows that are at steady-state. This identification is useful to select data sets for pa-
278 rameter identification with steady-state mathematical models. When processing large
279 amounts of data, this identification typically yields many data windows that are deemed
280 acceptable for parameter estimation. Taking similar data sets for the parameter fit gen-
281 erally leads to poor results because there is not enough data diversity to fit parameters
282 in nonlinear relationships. One example of this is that lack of temperature variabil-
283 ity in a reactor data does not allow activation energies to be identified because of the
284 co-linear relationship with the pre-exponential factor as shown in Table 4. If the tem-
285 perature data varies, a tighter confidence interval can be obtained for both E_a and k_0 .
286 The steady-state identification procedure shown in this work can be applied to an op-
287 timization problem with an objective to obtain the best limited number of diverse data
288 sets from a potential candidate pool.

289 Even though this technique is applied in time blocks, it can also be applied in a time
290 shifted approach to signal plant steadiness or unsteadiness on a continual and real-time
291 basis. For example, if a new sample is obtained every second, the past 3 to 5 time
292 constants could be used to determine the probability that the process is currently at
293 steady-state.

294 4. Conclusion

295 Presented in this work is a straightforward technique to effectively detect inter-
296 vals or windows of steady-state operation within continuous processes subject to noise.
297 This detection is critical in applications that rely on steady-state models for data rec-
298 onciliation, drift detection, and fault detection. The algorithm has minimal computing
299 requirements involving statistical estimates and has only two settings to specify i.e.,
300 the window-length and the Student-t critical value. Multivariate systems can be easily

301 handled by including several key process signals and adjusting the critical Student-t
302 statistic accordingly. Finally, the benefit of detecting windows of steady-state behavior
303 in a plant with multiple interacting major processing units for example can be useful
304 even by itself without executing on-line steady-state models for monitoring or opti-
305 mization.

306 **References**

- 307 [1] A. Mhamdi, W. Geffers, F. Flehmig, W. Marquardt, *Desalination and Water Re-*
308 *sources: Thermal Desalination Processes*, Vol. 1, EOLSS, 2010, Ch. On-line op-
309 *timization of MSF desalination plants*, pp. 136–162.
- 310 [2] W. Holly, R. Cook, C. Crowe, *Reconciliation of mass flow rate measurements in*
311 *a chemical extraction plant*, *Can. J. Chem. Eng.* 67 (1989) 595–601.
- 312 [3] R. Bethea, R. Rhinehart, *Applied Engineering Statistics*, Marcel Dekker Inc.,
313 *New York*, 1991.
- 314 [4] S. Narasimhan, C. Kao, R. Mah, *Detecting changes of steady states using mathe-*
315 *matical theory of evidence*, *AIChE J.* 33 (1987) 1930–1932.
- 316 [5] S. Cao, R. Rhinehart, *An efficient method for on-line identification of steady-*
317 *state*, *J. Proc. Cont.* 5 (1995) 363–374.
- 318 [6] M. Mansour, J. Ellis, *Methodology of on-line optimization applied to a chemical*
319 *reactor*, *Appl. Math. Modell.* 32 (2008) 170–184.
- 320 [7] M. Kim, S. H. Yoon, P. Domanski, W. Payne, *Design of a steady-state detector*
321 *for fault detection and diagnosis of a residential air conditioner*, *Int. J. Refrig.* 31
322 (2008) 790–799.
- 323 [8] M. Campos, H. Teixeira, F. Liporace, M. Gomes, *Challenges and problems with*
324 *advanced control and optimization technologies*, in: *7th IFAC Intern. Symp. Adv.*
325 *Contr. Chem. Proc.*, Istanbul, Turkey, 2009.
- 326 [9] G. Box, G. Jenkins, *Time-Series Analysis: Forecasting and Control*, Revised Edi-
327 *tion*, Holden Day, Oakland, California, 1976.
- 328 [10] J. Kelly, *Tuning digital pi controllers for minimal variance in manipulated input*
329 *moves applied to imbalanced systems with delay*, *Can. J. Chem. Eng.* 76 (1998)
330 967–974.
- 331 [11] S. Narasimhan, C. Jordache, *Data Reconciliation and Gross Error Detection: An*
332 *Intelligent Use of Process Data*, Gulf Publishing Company, Houston, TX, 2000.
- 333 [12] J. Ahrens, K. Kohrt, U. Dieter, *Algorithm no.599*, *ACM-Trans. Math. Soft.* 9
334 (1983) 255–257.
- 335 [13] J. Hedengren, T. Edgar, *Approximate nonlinear model predictive control with in*
336 *situ adaptive tabulation*, *Computers and Chemical Engineering* 32 (2008) 706–
337 714.