Antonis Kokossis, Michael C. Georgiadis, Efstratios N. Pistikopoulos (Eds.)
PROCEEDINGS OF THE 33rd European Symposium on Computer Aided Process Engineering (ESCAPE33), June 18-21, 2023, Athens, Greece

A simple two-parameter steady-state detection algorithm: Concept and experimental validation

Evren Mert Turan[a], Johannes Jäschke[a],

[a] *Department of Chemical Engineering, Norwegian University of Science and Technology, Sem Sælandsvei 4, 7491, Trondheim, Norway*
johannes.jaschke@ntnu.no

Abstract

Automatic detection of steady state periods is a necessary step for many tasks related to real time operation of a process. Based on the insight that the time series of a controlled process variable at steady state resembles a mean reverting process, we fit a first order auto-regressive model to a window of data and use the Dickey-Fuller test to test for this mean reverting property. We compare the proposed approach on two synthetic data sets and one experimental data set and find that the method performs well in comparison to methods in the literature. As the method is computationally inexpensive, only requires two parameters, and is interpretable, we suggest that this is an effective steady state detection tool.
Keywords: Steady state detection, Process Monitoring.

1. Introduction

The identification of steady state periods in industrial process operations is an essential initial step for many typical tasks, e.g. parameter estimation, abnormal event identification, and real time optimization. Incorrect identification of steady state can severely impact operation, e.g., updating a steady state model with dynamic data can lead to reduced profits and/or unstable operation. Although steady state can be reliably identified by human operators (in low-dimensional data), the automatic identification of steady state, especially with substantial amounts of online data in high dimensions, is still challenging. In this work we use the Dickey-Fuller test to identify steady state time windows, and demonstrate it's effectiveness in comparison to other methods in the literature based on a their perforance on experimental and synthetic data.

Most approaches in for steady state detection in the process engineering literature use 1) a statistical test, or 2) a classification based machine learning method. In this work we focus on the former due to their interpretability, and simple transfer between applications due to their relatively small amount of tuning parameters. Methods relying on statistical tests either tend to perform this test on characteristic values of the system or parameters of a model fitted to online data (Rhinehart, 2013). A key difference between such methods is in the construction of the statistical test, i.e. whether steady state is the null or alternate hypothesis. Although most methods have steady state as the null hypothesis, identification of steady state by the alternate hypothesis is a statistically stronger claim (Rhinehart, 2013).

In this work we fit an auto-regressive (AR) model and use the Dickey-Fuller test to test the null hypothesis that the process is transient. The method is simple to tune, having only two parameters: the size of the time window and the significance level used in the statistical test. In process engineering, the Dickey-Fuller test has been used in fault detection, however to the authors knowledge it has not been applied to steady state detection.

2. Background literature

We briefly outline three methods from the steady state detection literature that make use of a statistical test. In all these methods some quantity based on recent measurements is determined (e.g. a slope), and a statistical test is performed on this quantity (Rhinehart, 2013). Furthermore, in the first two methods we consider a rolling time-window of *N* recent measurements.

2.1 Method 1: Slope of a line

The simplest approach to steady state detection is to fit a linear model to the data:

$$y_k = mt_k + c \tag{1}$$

where y is a measured process variable, t_k is the k^{th} time point, m is a slope, and c the intercept. After fitting this line by linear least squares, one can check the null-hypothesis of steady state by testing if the slope is zero using a t-test at some significance level. However, this approach violates some assumptions, in particular the y variables at different time points are not independent, i.e. y_{k+1} depends on y_k, and y_{k-1} and so on.

2.2 Method 2: Kelly and Hedengren

In this approach, we test if the process can be described as being at some value, and subjected to independent, identically distributed white noise (Kelly and Hedengren, 2013). As before, we consider the linear model in equation 1. To estimate the parameters averages of the intercept, *c*, and gradient, *m*, are used. Note that if there is gradient is zero then the intercept is the mean value. If the data is sampled uniformly in time by Δt, then these are given by:

$$m = \frac{1}{N\Delta t}\sum_{k=2}^{N} y_k - y_{k-1} \tag{2}$$

$$c = \frac{1}{N}\left(\sum_{k}^{N} y_k - m\Delta t\right) \tag{3}$$

The standard deviation of the noise, σ_n, can then be estimated as:

$$\sigma_n = \sqrt{\frac{\sum_k^N (y_k - mt_k - c)^2}{n-2}} \tag{4}$$

Lastly, steady state identification is performed pointwise by checking if the deviation of the data point from the intercept is within a factor of the standard deviation of the noise. This cut-off point is given by a Student's critical value, t_{crit}, at some significance level and degree of freedom *N*:

$$s_k = \begin{cases} 1 & \text{if } |y_k - c| \leq t_{crit}\sigma_n \\ 0 & \text{else} \end{cases} \tag{5}$$

$$S = \frac{1}{N}\sum_{k=1}^{N} s_k \tag{6}$$

where S represents the fraction of time in the window "at steady state". If $S \geq 0.5$, then the entire window is at steady state.

2.3 Method 3: Cao and Rhineheart

This method relies on the idea that the ratio of two different estimates of the time series' variance at the same time point should be close to one when the process is at steady state (Cao & Rhinehart, 1995). The first estimate of the variance at t_k, $\sigma_{1,k}^2$, is given by the filtered squared difference between the process measurement, and a filtered value of the measurement, $y_{f,k}$. First the filtered measurement is calculated, and used to calculate a filter of the mean square deviation, v_k^2:

$$y_{f,k} = \lambda_1 y_k + (1 - \lambda_1) y_{f,k-1} \tag{7}$$

$$v_k^2 = \lambda_2 \left(y_k - y_{f,k}\right)^2 + (1 - \lambda_2) v_k^2 \tag{8}$$

Assuming the process is at steady state and the measurements are uncorrelated, then the variance, $\sigma_{1,k}^2$, can be directly estimated as (Cao & Rhinehart, 1995):

$$\sigma_{1,k}^2 = \frac{2 - \lambda_1}{2} v_k^2 \tag{9}$$

The second estimate of the variance, $\sigma_{2,k}^2$, is given by filtering the difference between two consecutive measurements:

$$2\sigma_{2,k}^2 = \lambda_3 (y_k - y_{k-1})^2 + (1 - \lambda_3) 2\sigma_{2,k-1}^2 \tag{10}$$

Lastly, the ratio of these two variances are taken to calculate an R statistic:

$$R_k = \frac{\sigma_{1,k}^2}{\sigma_{2,k}^2} \tag{11}$$

The null hypothesis is that the process is steady, corresponding to a ratio close to 1. If R_k is greater than some critical value, R_{crit} then the null hypothesis is rejected.

3. Proposed method

The proposed method is based on the insight that the time series of a controlled process variable at steady state, subject to stochastic disturbances, resembles that of a *mean reverting process*, i.e. it tends to some mean point despite the stochasticity of the system. In the example of a controlled process unit, this would be due to the relevant controllers rejecting disturbances, or the system settling to a new operation point.

The first step in the appraoch is to find the mean value in the time window, μ, and subtract this from each data point:

$$\tilde{y}_k = y_k - \mu \tag{12}$$

Then we consider the first order auto-regressive (AR) model:

$$\tilde{y}_{k+1} = p\tilde{y}_k + n_k \tag{13}$$

where $\tilde{y}_k$ is the transformed measurement at time t_k, n_k is some random variable with mean zero and finite variance, and p is a variable estimated by linear least squares on the data in the time window. A value of $|p| < 1$ means that the process is mean reverting, as as the absolute disturbance from the mean will deterministically decrease at each iteration. After estimating $\hat{p}$, we then perform a one-sided confidence test on the null-hypothesis that $|p| = 1$, i.e. the process is in a transient state, by calculating the test statistic τ:

$$\tau = (\hat{p} - 1)\sqrt{\frac{(N-2)\sum_{k=2}^{N} \tilde{y}_k^2}{\sum_{k=2}^{N}(\tilde{y}_k - \hat{p}\,\tilde{y}_{k-1})^2}} \tag{14}$$

If the test statistic is less than the critical value from the Dickey-Fuller distribution then the null hypothesis is rejected, i.e. the process is at steady state. The critical values can be found from response surfaces fit to the distribution (MacKinnon 2010). The asymptotic results of this test do not require assumptions of normality or homoscedasticity (MacKinnon 2010). Lastly, note that if the process variable is controlled off-set free to some set-point, then one could use the set-point instead of calculating the mean value.

4. Results

We compare the proposed method (DF) against the line slope, Cao & Rhinehart (CR, 1995) and Kelly & Hedengren (KH, 2013) methods on two synthetic data sets and one experimental data set.. For brevity these acronyms are used throughout this section. For these data sets we label steady state and transient periods based on 1) knowledge of when the system inputs are changing and 2) inspection of the data in the time windows. For the methods with time windows and significance level as hyper-parameters we use a time window of 30 seconds, and significance level of 95%. For the method of Cao & Rhinehart (1995) we use the recommended values of $\lambda_1 = 0.2, \lambda_2 = \lambda_3 = 0.1,\ R_{crit} = 2$.

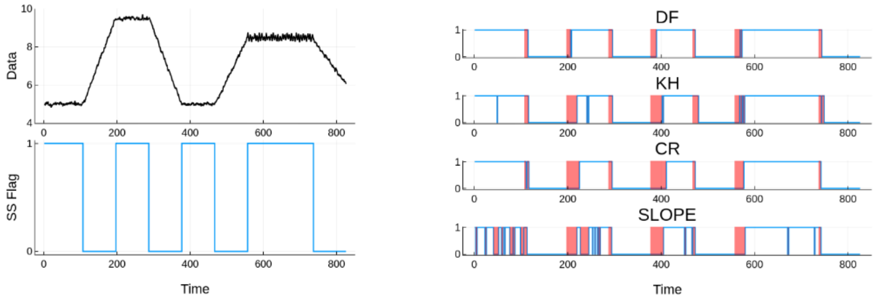
4.1 Low level of normally distributed noise

We first consider a synthetic example with a low level of Gaussian noise. This is an easier task than the others and therefore serves to give a baseline of the methods performance. The results are summarized in Figure 1. From a visual inspection, the proposed DF method performs the most consistently, with the worst performance by the slope method. For all the methods the primary source of error is due to the delay that occurs when the system goes from transient to steady state (around 180s, 360s, and 540s in Figure 1). In contrast, the transition from steady to transient state is captured without a significant delay. The transition to steady state is harder to detect due to the presence of the transient in the time window, as this heavily influences the fit of the models and the filtered variance. Note that if the DF method is used with a set-point instead of the mean then this delay would be reduced, i.e. the delay for this approach comes from taking the mean of the data.

4.2 Moderate level of t-distributed noise

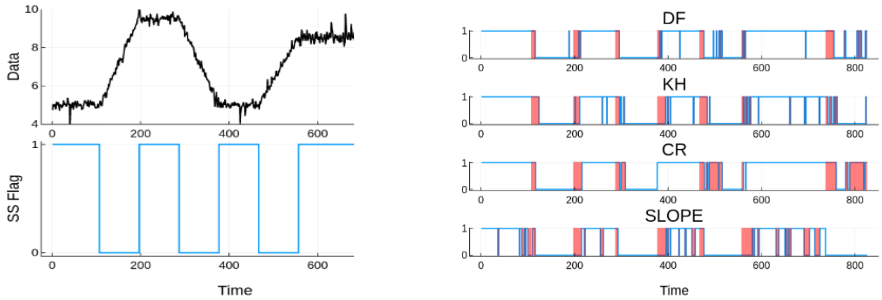
In this comparison we consider the same process as in section 4.1, but now use t-distributed noise, with a larger variance, instead of Gaussian noise. The t-distribution has heavier tails than the normal distribution leading to a high probability of "high-magnitude" noise, leading to a more challenging classification. The experiment is summarized in Figure 2, with "spikes" due to change in distribution clearly shown in Figure 2a. The aim of using this distribution is 1) to test the methods against higher noise levels, 2) to test against non-normally distributed noise.

Apart from the general decrease in performance, the clearest difference when comparing Figure 2b vs 1b is that DF, KH, and CR have a higher rate of false identification of steady state. This does not occur with slope method, as this approach is extremely sensitive to noise. As in section 4.1 the DF method has the best performance.



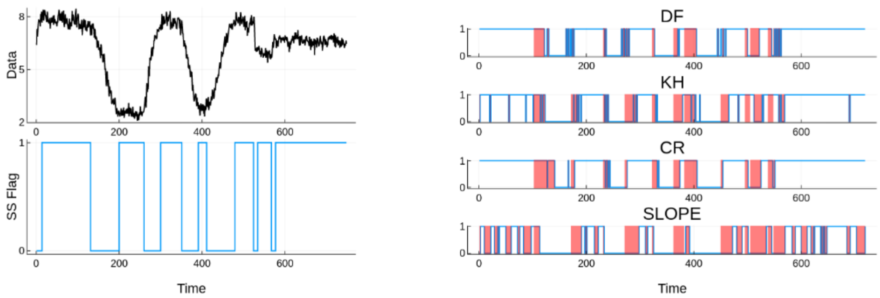
a b

Figure 1: Synthetic data with low levels of Gaussian noise and identified steady state periods (A), and steady state predictions (B). In A and B, a steady state flag of 1 indicates steady state. The shaded red regions in B indicate a mismatch between the predicted and true system state



a b

Figure 2: Synthetic data with moderate levels of t-distributed noise and steady state periods (A), and steady state predictions (B).



a b

Figure 3: Experimental data with steady state periods (A), and steady state predictions (B)

4.3 Experimental data

Lastly, we consider the steady state algorithms applied to real experimental measurements from a lab rig with the experiment summarized in Figure 3, and Table 1.

Table 1: Summary statistics of the steady state detection algorithms applied to the experimental data, shown in Figure 3. The bolded entries indicate the best method for a statistic

STATISTIC	DF	KH	CR	SLOPE
PRECISION	0.85	0.83	0.81	**0.86**
RECALL	**0.90**	0.78	0.89	0.44
F1 SCORE	**0.87**	0.81	0.85	0.58
ϕ COEFFICIENT	**0.60**	0.46	0.51	0.29

This is a real system with a relatively large amount of noise, leading to worse performance compared to the synthetic results by all the methods. In addition there are two step changes between 500 and 600 seconds, which results poorer performance in this section, as shown in Figure 3. A positive aspect of the CR and DF methods are that they are more consistent as they have much less false flags within the SS and transient periods.

Summary statistics of the methods on this data set are shown in Table 1 and serves as quantitative evidence of the visually better performance of the DF method in Figure 3. Precision is the fraction of correct SS predictions over all SS predictions, while recall is the fraction of correct SS predictions over all true SS periods. The F1 score is the harmonic mean of precision and recall. Lastly the ϕ coefficient is a balanced metric that requires reliable performance in both SS and TS prediction. In general, the proposed DF method performs the best, except for the precision metric where the simple line slope method performs similarly due to rarely identifying steady state, as shown in Figure 3.

5. Conclusions

In conclusion, we present the application of the Dickey-Fuller test for use in steady state detection and compare it with other methods in the literature on a range of examples. Based on summary statistics of these methods, the Dickey-Fuller test performs the best overall. The method is simple to implement and only requires two-hyperparameters. Further work could extend the approach to multivariet system using approaches suggested in the literature (Rhinehart, 2013, Kelly and Hedengren, 2013).

Acknowledgments

The authors acknowledge the support of the Norwegian Research Council through the AutoPRO project.

References

S. Cao and R.R. Rhinehart, 1995, An efficient method for on-line identification of steady state, Journal of Process Control, 5(6), pp.363-374.

D.A. Dickey and W.A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root. Journal of the American statistical association, 74(366a), pp.427-431.

J.D. Kelly, and J.D. Hedengren, 2013, A steady-state detection (SSD) algorithm to detect non-stationary drifts in processes, Journal of Process Control, 23(3), pp.326-331.

J.G. MacKinnon, 2010, Critical values for cointegration tests, Queen's Economics Department Working Paper, No. 1227.

R.R. Rhinehart, 2013, Automated steady and transient state identification in noisy processes, American Control Conference, pp.4477-4493.