Short-Term Transition Impact on Change-Point Detection and Monitoring Existing State

Turkan K. Gardenier[1]
1Teka Trends, Inc. 1733 Kirby Rd., #1408, McLean, VA 22101

Abstract
Tracing short-term changes in monitored data is an important component of analytic modeling for interventions. Identifying change-point in time-series involves sensitivity to the possible, and often existing, autocorrelation between adjacent observations. This paper presents an approach which introduces the feasibility to evaluate successive data along a time series. A five-category interval with mid-range at state "zero" and two on each side are used, reminiscent of "control bands." An application to petroleum stock monitoring including motor gasoline and crude oil is presented to graphically depict differences between observed data and pre-established bands. The indices developed by the author, Cumulative Transitional State Score (CTSS) and Transitional Score (TS) are superimposed on the charts and are compared over time in order to detect and yield direction of change. A two-player game, with tabulation instructions and poem introduces the concept of successive transitions and their impact on signaling emerging change.

Key Words: Monitoring, Change Point Identification, Cumulative Transitional State Score (CTSS), Transitional Score (TS), Step Change, Slope Change, Control Charts, Cumulative Sums (CUSUM)

1. Comparative Monitoring Principles

1.1 Control Charts and Cumulative Sums (CUSUM)

Control charts were initially used in industry in detecting changes in the overall mean value of the process over time. The most commonly used control chart is the Shewhart chart. An interval is established representing a buffer zone of uncertainty around the mean value. Data are plotted over time along these boundary zones. Decision rules are established showing how many successive observations outside these boundary zones need to be observed in order to conclude that the process is out of control, or that the previously observed steady state has changed. The decision to accept or reject steady state conditions is based upon the assumption that successive observations are identically and independently distributed (i.i.d.) and have common variances. Normality of the distribution is also assumed.

Two parameters, h and k are used to determine if the overall process has changed. As described in Johnson and Leone (1962), Lucas (1976) and Page

(1961). An alternative average is accepted if the established boundaries are crossed.

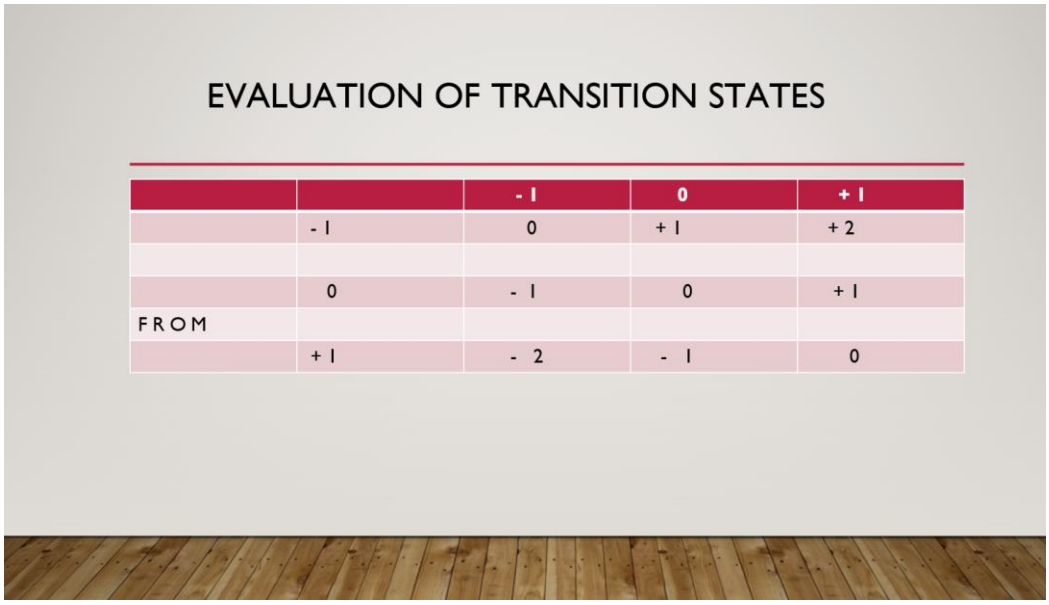
1.2. The Cumulative Transitional State Score (CTSS) Incorporating Short-Term Changes

The Cumulative Transitional State Score (CTSS) was developed as a measure for trend evaluation and monitoring (Gardenier 1979, 1984, 1997, 2019). While CUSUMS cumulate the sum of deviations from a long-term average, CTSS uses a cumulative soring scheme applied to transitions between successive observations. CTSS defines "states" as regions or partitions, reminiscent of zones in Shewhart control charts. One may use three zones, akin to 95% tolerance limits in control charts, or multiple zones, as in five-line control chart procedures.

Table 1 shows possibilities for assigning a score for change in adjacent observations. FROM, in rows, refers to the prior observation while TO, in columns, refers to the nexr successive observation. The transition FROM to TO generates the TRANSITION index, or transitional score (TS) by evaluating boundary crossings. If there is no change, observations remain within -1, zero or +1 boundaries, as shown in the diagonals of the 3x3 matrix, the TS score is zero. If observations increase to the next level, from -1 to zero,or zero to +1, the TS score is +1. If they cross two boundaries by increasing from -1 to +1, the TS score is +2. In the same vein, a decrease from zero to -1 yields a TS score of -1; a decrease from +1 to -1 yields -2 because two boundaries are crossed.

TS scores are added or cumulated over each successive observation. This process generates the Cumulative Transitional State Score (CTSS).. .

Table 1: Framework for Assigning 0, -1, and +1 to Successive Observations



2. Illustrative Application to Petroleum Stocks

2.1 Scope and Content of Database

2. 1.1 Motor Gasoline and Crude Oil Stocks (millions of barrels)
Data were based on computerized Weekly Petroleum Status Report (1983)of Energy Information Administration, DOE (EIA-0208)83/45. Monthly data for each of two series, Motor Gasoline and Crude Oil were used in the present analyses.

2.1.2 Seasonal Adjustment Bureau of the Census X-11 method
Each record was seasonally adjusted as reported in Technical paper 15 of the US Bureau of the Census (Shishkin et. al. 1967), thus generating the stock bands in the present report.

2.2 Application of CTSS and TS Scoring

Figure 1 and supplement for four prior years portray monthly data for U.S. stocks of Motor Gasoline (top figure) and Crude Oil (bottom figure) during 1980-1982 in millions of barrels. Data show observations and four lines, creating three intervals, which are BANDS of expected valuers. The middle interval may be denoted asa average or zero, the interval above +1, the one below -1.

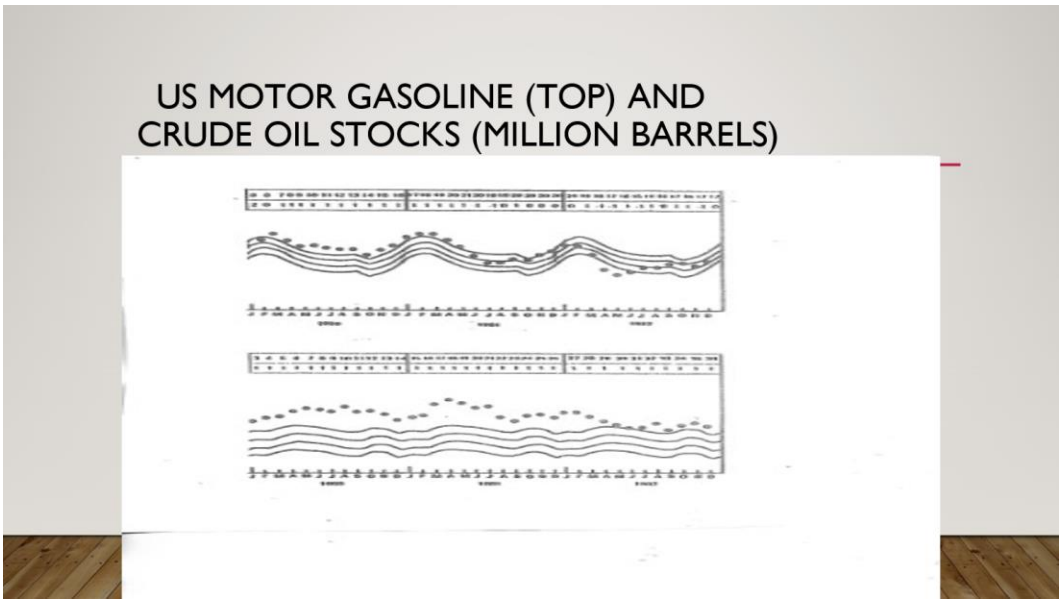
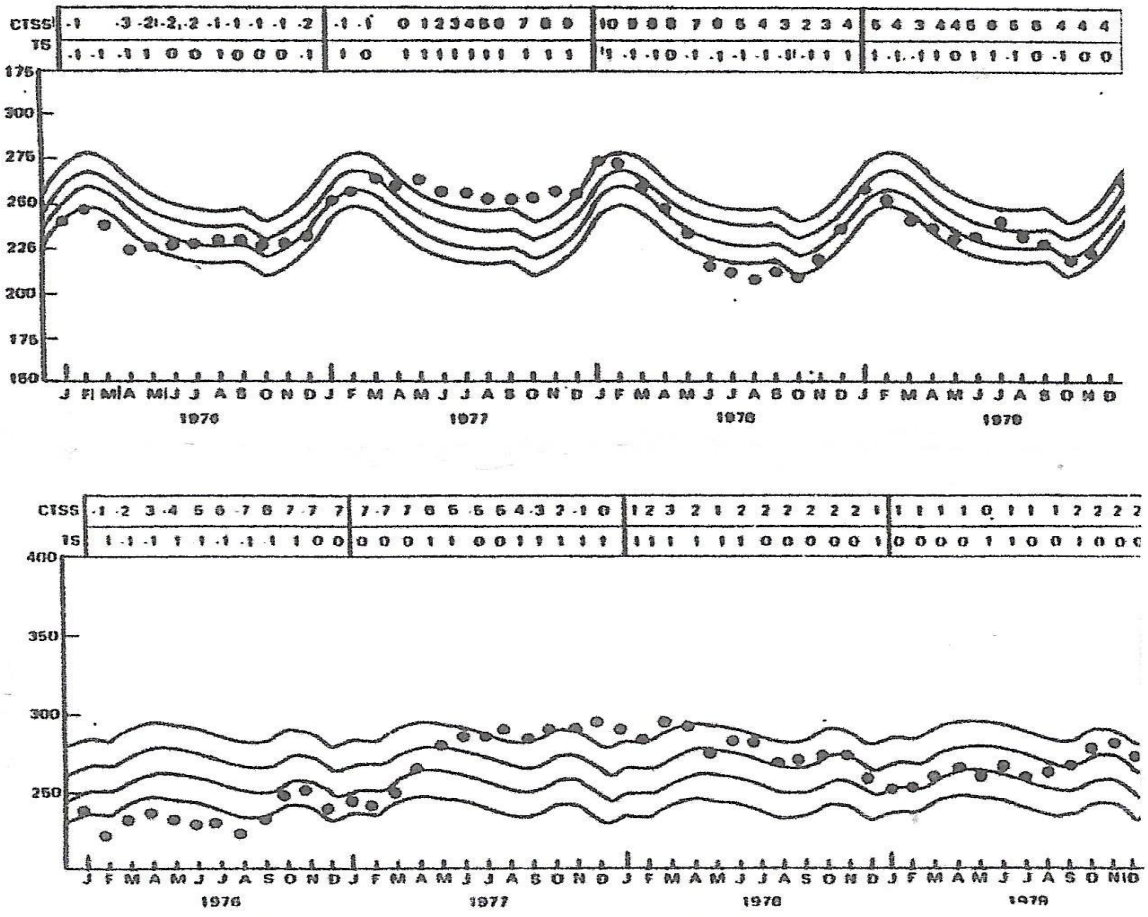


Figure 1: U.S. Petroleum Supply of Motor Gasoline and Crude Oil over 3 Years

Scanning the data shows that Motor Gasoline data are mostly within the bands, but data for Crude Oil are mostly above the bands. A more thorough examination of the data for the prior four years (1976-1979)7is shown below. The top figure shows data for Motor Gasoline, below are data for Crude Oil. Motor Gasoline data appear to follow seasonality

expectations.



During the early time period Crude Oil Stocks started berlow boundary lines, then had a steady upward ncrease reaching above boundary lines, then showed a decrease followed by another increase.

CTSS and TS are useful in tracking such changes and making preparing for stock control decisions during the following three years, which we already observed. The TS and CTSS values appear as the first two lines. Differences in CTSS scores between two time periods are informative as well.

2.3 Graphical Representation of Trend Shift

Two types of time-oriented time-oriented plots are shown in the top and middle sections of Figure 2 A straight line has been fitted to rising and/or decreasing successive CTSS scores in the top section, making us aware of emerging trends. Coordinating with personnel involved in stock control inputs may be prepared for supply chain management issues such as transportation, processing and storage.

The middle section displays the sequence of "run"s of TS scores, zero, 1 or -1, over time. Sequence of zeros implies no change or stablility, +1s indictes rising trend in petroleum stocks, -1 indicates a continued decrease in stocks. The information gleaned is much more informative than the plot of actual dta in the lower section.

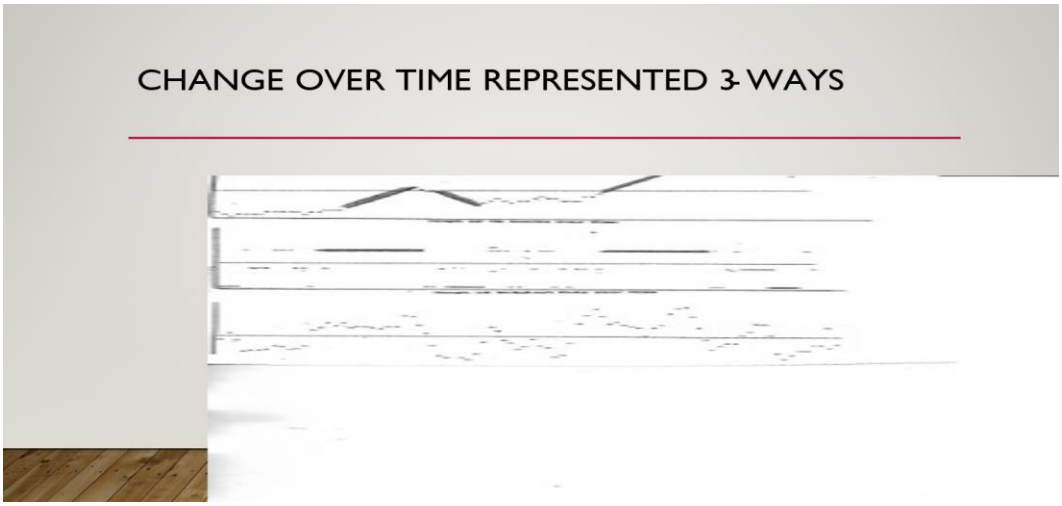
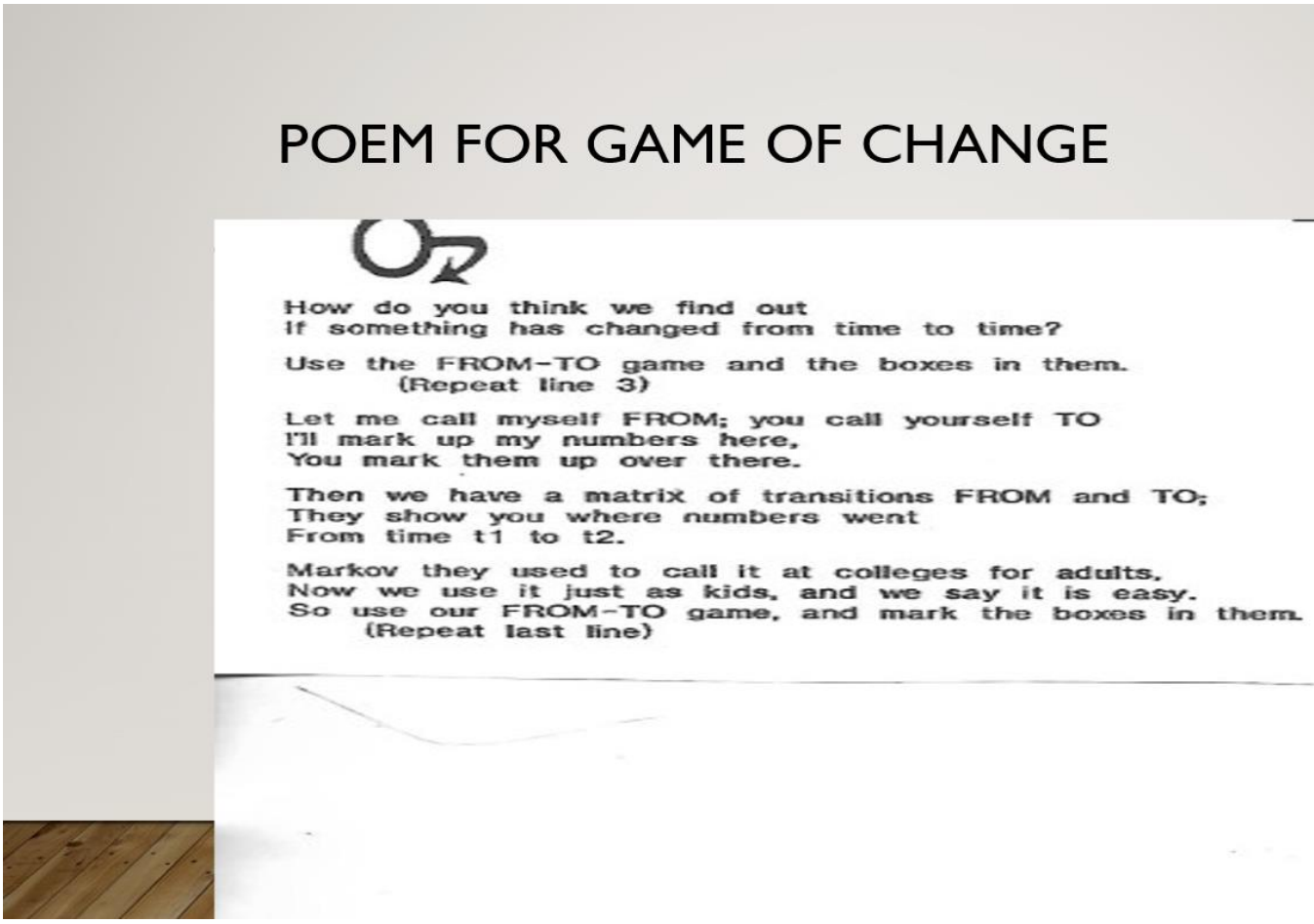


Figure 2: Illustration of Trend Patterns Emerging from 0/-1/+1 Assignments
CTSS in Top Section; TS in Section Below Represented as Sequences

3. Two-Person Game Based on CTSS and TS

In Gardenier (1997) the principles underlying the Cumulative Transitional State Score and the concept of successive transitions were introduced to school-age children using games and songs. In a two-player game digits, representing adjacent observations, are called out. The players tabulate them into transition categories FROM and TO, which is the basis of the present monitoring scheme.



.

Figure 3. Text to Accompany Interactive 2-Person Game

.

Acknowledgements

The author wishes to acknowledge the encouragement and support of two deceased members of the statistical community: Yvonne M M. Bishop,, Energy Information Administration, U.S. Department of Energy, Washington, D.C. and Gladys Reynolds of the Centers for Disease Control, Atlanta, Georgia..

References

Gardenier TK. Profiling for efficacy in long-range clinical trials. 1979. Advanced Medical Systems: An Assessment of the Contribution. Chicago, IL: Yearbook Meds. Publ., 95-102.

Gardenier TK. CTSS/CUSUM applications to interverntion analysis. 1984. In JSM Proceedings, Survey Research Methods Section, Alexandria, VA: American Sttiatical Association, 777-781.

Gardenier T K. Time: In Gardenier Math-Stat Serries. 1997. Vienna, VA. Pro-File Inst.

Gardenier TK. Step-function approach to time series:an air quality application., International J. Applied Science and Technology 2013; 3(8): 15-20.

Gardenier TK. Time Trends. 2019. McLean, VA. Teka Trends, Inc.
Gardenier TK, Gardenier JS. Delving into megadata: evolving challenges 2013. In JSM Proceedings, Statistical Learning and Data Mining Section, Alexandria, VA: American Statistical Asasociation, 324-333.
Gardenier TK, Gardenier JS. Trinary Natire of Big Data. 2018. McLean, VA. Teka Trends, Inc.
Hall DL, Gardenier TK, Slavich AL. Intervention adjustment of data of the joint petroleum reporting system. 1984. Final Report of the US Department of Energy Contract DE-AC06-76RLO-1830. Richland, WA: Battelle Pacific Northwest Laboratories.
Johnson NL, Leone FC. Cumulaive sum control charts: mathematical principles applied to their contribution and use. Industrial Quality Control 1962; 15-21.
Lucas JM. The design and use of V-Mask control schemes. J Quality Technol 1976; 8:1-12..
Page ES. Cumularive sum chartse. Technometrics 1961; 3-1-9.

.