Jurnal Matematika Vol. 16, No.1, Juni 2026, pp. 42-55 ISSN: 1693-1394
Article DOI: 10.24843/JMAT.2026.v16.i01.p198

Hotelling's T² Control Chart for Monitoring Correlated Wine Quality in Small-Scale Production

Ayenigba[1*] ,Deborah Ayomide Adegboye [2], Dorcas Omobola Folarin[3], James Serumun Ivande[4]

[1,2,3] *Department of Mathematical Sciences, Ajayi Crowther University Oyo, Nigeria.*
[4]*Department of Statistics, Joseph Sarwuan Tarka University, Makurdi, Nigeria*

*Corresponding Author. Email: aa.ayenigba@acu.edu.ng

ABSTRACT

This research article investigates how effective Hotelling's T^2 control charts are in monitoring the quality of correlated characteristics in small-scale red wine production. A purposive sample containing 50 observations was taken from the UCI Machine Learning Repository and from Kaggle, which focuses on four physicochemical properties, including fixed acidity, volatile acidity, alcohol content, and residual sugar. Upon preliminary analysis, it was confirmed that there are dependencies between variables, as the strong positive correlation ($r = 0.668$) between alcohol and residual sugar, and a moderate negative correlation ($r = -0.435$) between fixed and volatile acidity warrants the use of a multivariate monitoring framework. Using the T^2 statistic, in a Phase I context, four out-of-control observations have been identified ($samples\ 2, 4, 8, and\ 13$). In contrast, parallel univariate Shewhart $\bar{X}$ charts only identified two out-of-control observations. To check sensitivity, a coordinated mean shift of 1.5 σ (standard deviations) was introduced over the last 10 observations. The T^2 chart showed a 40% detection rate at this level, while the univariate charts indicated no signal out of control. Due to the sample-to-variable ratio of $12.5:1$, the covariance matrix estimate did not exhibit a clearly favourable situation in favor of one method over the. In essence, the Hotelling's T^2 chart exhibits improved sensitivity to joint mean-vector deviations, effective control of familywise Type I error, and diagnostic decomposition in multivariate quality control applications for the purpose of process improvement.

Keywords: Hotelling's T²; Multivariate control chart; Wine quality; Statistical process control; Mahalanobis distance; Small-scale production.

I. Introduction

The quality of a wine is affected by the interaction of a high number of physicochemical variables, which evolve in a simultaneous manner in a chemical context. The properties of fixed acidity, volatile acidity, alcohol, and residual sugar vary in ways influenced at the same time by grape variety, fermentation management, and post-fermentation handling [13]. Monitoring these features in isolation, as classical univariate control charting prescribes, yields an inherently fragmented picture of process behavior. A joint shift in the mean vector of correlated variables may not cross any individual 3σ control limit, even though the multivariate system has meaningfully drifted away from in-control [16]. The existence of this detection gap has a real economic cost in lower production environments such as artisanal and boutique wineries where each production batch is worth considerable money and costly deviations inferred but undetected systemically cause intolerable product variation.

Small-scale producers have specific operational constraints that inhibit effective quality assurance, including small batch sizes, limited analytical resources, and staff mainly experts

in viticulture rather than statistics. However, due to protected designation of origin frameworks and supermarket private label quality requirements [20], there has been a much greater competitive and regulatory drive for demonstrable batch-level consistency. Using multivariate statistical process control (MSPC), small producers can satisfy these demands without the need for large-scale infrastructure. It is a rigorous and computable technique. The Hotelling T^2 statistic collapses the p-dimensional deviation of an observation from its in-control mean vector into one scalar using the Mahalanobis distance. Also, it takes into account the correlation structure of the quality attributes directly. Moreover, it produces a single nominal false alarm rate irrespective of the number of variables being monitored simultaneously.

The multivariate monitoring capabilities in similar small-batch food productions have been practically observed with applications of Hotelling's T^2 and MSPC. The production of sachet water [2], manufacture of soap [1], processing of soybean meal [5], and quality control of crude palm oil [6]. In the above cases, multivariate methods have been consistently more effective than corresponding univariate methods. In detecting the simultaneous shifting of the process variables. Wine quality pertains to the physicochemical character of wine. This quality has a strong correlation between its many variables. This was documented in a comprehensive manner by [13]. Thus, the present study uses covariance-aware Mahalanobi's theoretical framework as the best fit for this purpose. [18] showed, from different food and beverage matrices, that multivariate statistical methods reveal compound quality interactions that are completely unobservable to univariate analysis; this is directly applicable to fermentation-derived quality interactions. Even with the above proof, there is very little peer-reviewed literature on systematic MSPC deployment in artisanal wine production [20]. The present study tackles this gap directly.

Recent advances in statistical process monitoring have expanded significantly the practitioner toolbox. [12] report on how the evolution from classical Shewhart-type charts to memory-type and data-driven monitoring systems has been traced, with the superior detection performance documented when correlation between variables are modeled. According to [20], practitioner unfamiliarity (rather than a methodological limitation) is the most significant barrier to the adoption of multi-systems principal components (MSPC) in the food industry. Given that, it would be useful to have carefully documented applied studies of the sort we showcase here to close this gap. As noted in [7], the real-time multivariate monitoring of correlated performance parameters in collaborative manufacturing systems is beneficial. The result showed that monitoring in collaboration or jointly would reduce diagnostic time above that of univariate monitoring in sequential order. In this way, the same advantage would work in the study of wine batch monitoring.

This research makes three interrelated research contributions. An important objective of the study on wine physicochemical characteristics is to present the correlation structure among the four most important physicochemical variables and to construct a Hotelling's T^2 chart that reflects this structure so that we can establish a Phase I baseline performance under realistic small-batch conditions. In the second contribution, the multivariate chart is benchmarked against four univariate Shewhart $\bar{X}$ charts with parallel structures through observational comparison and simulation of a 1.5σ process-shift under controlled conditions so that the relative detection sensitivity and familywise error control can be quantified. In the third instance, the study will assess whether a 12.5:1 sample-to-variable ratio, using a dataset of 50 samples, is sufficient for stable Phase I covariance estimation in a feasible operational context.

This will be carried out in accordance with the guidelines outlined in references [14, 16]. The research is organized such that Section 2 describes the data, variables, and analytical methods, including a flow diagram of the methodology; Section 3 presents and interprets the findings; and Section 4 concludes with directions for further research.

II. Material And Methods

2.1 Rationale for Multivariate Monitoring

Conventional univariate Shewhart $\bar{X}$ charts monitor individual quality attributes independently. However, monitoring multiple correlated attributes at the same time may lead to a large false alarm rate. When lot sizes are large enough to assure an adequate number of defective parts, p independent charts can be run at the individual nominal false alarm rate α. However, this generates an overall or familywise Type I error of approximately $1 - (1 - \alpha)^{p}$. For p = 4, this gives an overall type I false-alarm rate of approximately 1.07% when $\alpha = 0.0027$. This is nearly four times the nominal rate α in this example. Moreover, the nonnegligible false alarm rate generates a lot of false corrective-action signals. Over time, this robs operators of trust and wastes resources [14, 17]. To resolve this issue, the T^2 statistic proposed by Hotelling computes the Mahalanobis distance of each observation from the mean vector in-control, wherein the contribution of the variables is weighted by S^{-1}, the precision matrix that encodes the full covariance structure. This generates a single scalar chart statistic with an upper control limit (UCL) set by the F-distribution, which maintains the nominal α regardless of p. The theoretical and empirical superiority of this approach over joint univariate schemes has been long-established in the statistical process control literature [8, 16]; this study provides a direct empirical demonstration in the under-researched context of artisanal wine production.

The T^2 approach has diagnostic decomposition as practical motivation. When the multivariate signal is raised, one can partition the out-of-control T^2 statistic according to the Mason–Tracy–Young (MAY) method described in [8], into unconditional contribution and contribution due to each variable conditional on the rest, for finding the specific variable or combination of variables that caused the exceedance. This ability to diagnose issues, which is absent from a set of individual univariate charts, is particularly beneficial for production operations not running on a large scale since diagnosis and correction must take place quickly so that subsequent batches do not incur losses.

2.2Data Source

Data used in the study is a publicly available red wine quality dataset at UCI Machine learning repository and provided in Kaggle (https://www.kaggle.com/datasets/uciml/red-wine-quality). The original dataset [13] of Minho Verde red wine with 1,599 observations and 11 physicochemical input variables plus one expert sensory quality rating was used. A sample of 50 observations was selected purposefully to represent a plausible scenario of small batch artisanal production keeping the four most relevant variables for fermentation quality and end-product stability. To enable controlled head-to-head comparison of multivariate versus univariate detection sensitivity under a known process deviation, a second dataset was simulated by introducing a coordinated 1.5σ shift in Fixed Acidity and Alcohol into samples 41 through 50, following the shift-magnitude convention recommended in [16].

2.3Variable

The study considers four continuous physicochemical variables and an indexing variable for monitoring the process. In order to increase production during fermentation, the fixed acidity reflects a potential imbalance of the structure in the wine. Whereas volatile acidity contains acetic acid. This may lead to spoilage of wine. The alcohol level (AL, % vol.) of a wine is a reflection of how far fermentation has gone. It has implications for sensory quality and stability. Residual sugar (RS, g/dm2) represents sugars that were not fermented during winemaking. RS gives the wine sweetness; however, RS interacts with acidity and alcohol, which can remain in the final wine product. A sample order and monitoring within the multivariate framework are maintained by an observation index (1-50).

2.4 Methodology Framework

According to the Phase I/ Phase II monitoring paradigm [14, 17], the analytical procedure was arranged in eight stages. As can be seen in flowchart, the entire sequence begins with stage 1, which performs exploratory data analysis (descriptive statistics, run charts, coefficient of variation). Stage 2 measures variable-to-variable correlations from the Pearson matrix and scatter-plot matrix. In Stage 3 multivariate normality is assessed using Mardia's skewness and kurtosis tests. Furthermore, chi-square Q-Q plots are also produced. If normality is rejected, Box-Cox transformation is applied. Stage 4 carries out the estimation of Phase I parameters and the construction of T² charts. Assignable-cause observations are iteratively removed on a time-sequence basis. Phase II future monitoring with finalized control limits is employed in Stage 5. Stage 6 builds 4 univariate Shewhart $\bar{X}$ charts for reference. Stage 7 executes the simulation of the 1.5σ shift. An MTY diagnostic decomposition of out-of-control signals is performed by Stage 8.

2.5Sample Size and Covariance Estimation Feasibility

The sample size m must strictly exceed p in order for the p × p sample covariance matrix S to be non-singular for reliable Hotelling T² monitoring. The sample to variable ratio is 12.5 to 1 with $m = 50 \; and \; p = 4$. [14] suggests that in order to obtain stable estimation of covariance, the value of m/p should be at least 10; this was confirmed through simulation by [16] for several covariance structures, which confirmed that performance of the T² chart will be reliable when its value is above 10. This $12.5{:}\,1$ ratio obtained here shows that this evidence supports the application of our framework for settings with small-scale production. An emphasis on smaller dimensionality settings distinguishes this application from high-dimensional settings where penalised covariance estimators [8] would be required.

2.6Hotelling's T² Statistic and Control Limits

Using Hotelling [11] and the Phase I/II formalization in [14, 17], let the observation vector at time i be:

$$X_i = \begin{bmatrix} X_{i,FA} \\ X_{i,VA} \\ X_{i,AL} \\ X_{i,RS} \end{bmatrix} \sim N_4(\mu, \sum)$$

where

$$\mu = \begin{bmatrix} \mu_{FA} \\ \mu_{VA} \\ \mu_{AL} \\ \mu_{RS} \end{bmatrix}$$

and $\sum$ is the 4×4 covariance matrix capturing cross- correlations

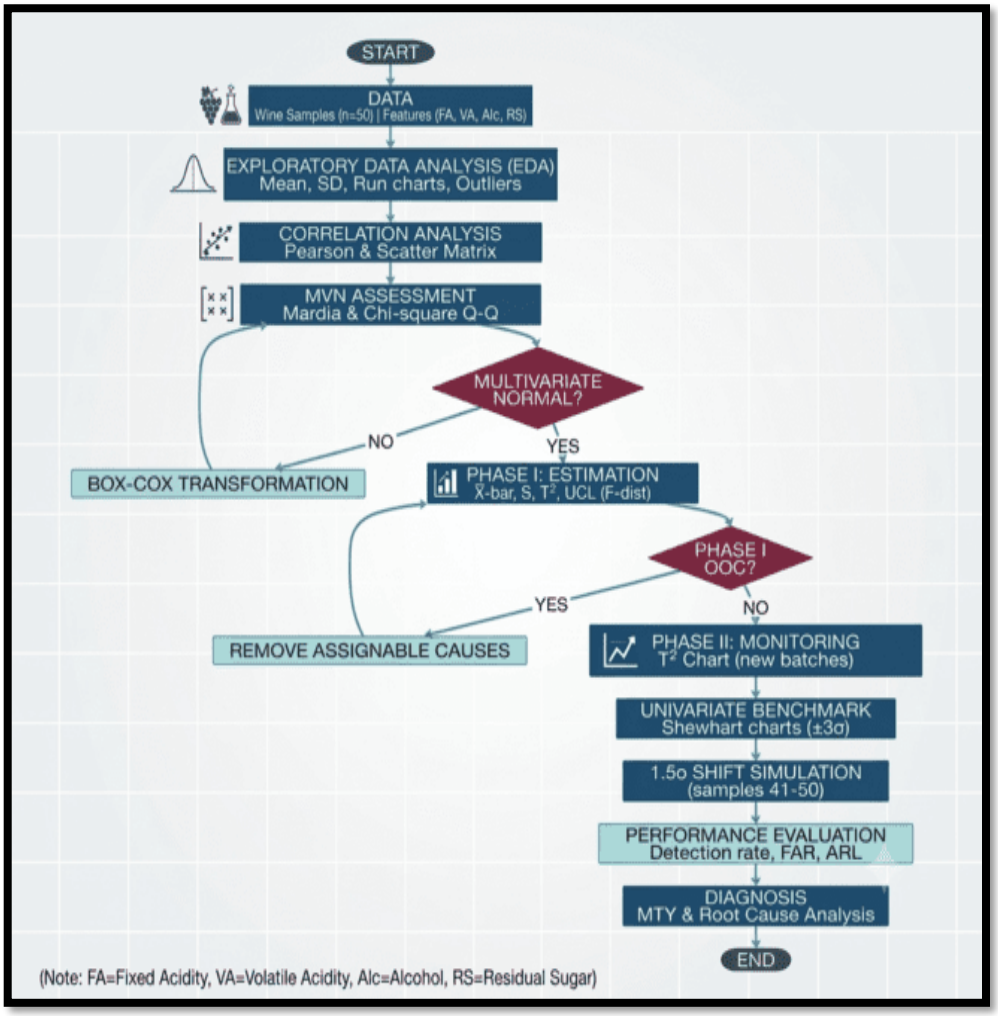

Figure 1. Methodology Flowchart – Hotelling's T² Multivariate Statistical Process Monitoring Framework

Phase I

For m subgroups of size n, subgroup mean vector is given by

$$\overline{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij} \qquad (1)$$

and overall mean vector

$$\overline{\overline{X}} = \frac{1}{m}\sum_{i=1}^{m}\overline{X}_i \tag{2}$$

$$S = \frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(X_{ij} - \bar{X}_i\right)\left(X_{ij} - \bar{X}_i\right)^T \tag{3}$$

The Hotelling Statistic

$$T_i^2 = n\left(\overline{X}_i - \overline{\overline{X}}\right) S^{-1}\left(\overline{X}_i - \overline{\overline{X}}\right) \tag{4}$$

which follows a scaled Beta distribution in Phase I. The exact Phase I UCL, derived from the F-distribution result formalized in [14, 16], is:

$$UCL^I = \frac{p(m-1)(n-1)}{mn - m - p + 1} F_{\alpha,p,mn-m-p+1} \tag{5}$$

where $p = 4, m = 50,$ and $\alpha = 0.05,$ a signal will be raised when $T_i^2 >$ UCL. Following the MTY decomposition methodology described in [8], the system resorts to the partitioning of the out-of-control signal T_i^2. The partitioning of T_i^2 into unconditional contributions $d^2(j) = \frac{(x_{ij} - \bar{x}_j)^2}{s_{jj}}$ and conditional contributions for every variable j allows the agent to identify root causes.

2.7Univariate Shewhart $\bar{X}$ Charts

To compare detection performance, individual Shewhart $\bar{X}$ charts were constructed for each quality variable. Assuming $X_{i,j,k} \sim N(\mu_{k,}\sigma_k^2)$

where $k \in \{FA, VA, AL, RS\}$

For sub grouped data:

$$UCL_K = \overline{X}_k + 3\frac{s_k}{\sqrt{n}} \tag{6}$$

$$CL_K = \overline{X}_k \tag{7}$$

$$UCL_K = \overline{X}_k - 3\frac{s_k}{\sqrt{n}} \tag{8}$$

We compared detection outcomes derived from the union of all four charts with T^2 signals under baseline and simulated-shift conditions. Using an $\alpha = 0.0027$ per chart, the familywise false alarm rate for four simultaneous independent charts is about $1 - (1 - 0.0027)^4 \approx 0.0107$. That is nearly four times the maintained 0.0027 for the T^2 chart [14]. This discrepancy encourages a simultaneous multivariate approach for statistical reasons as well as practical reasons.

III. Results And Discussion

3.1 Descriptive Statistics

Table 1. Descriptive Statistics of Wine Quality Variables (n = 50)

Variable	Mean	Std Dev	Min	Max	CV (%)
Fixed Acidity	7.486	0.675	5.60	11.20	9.02
Volatile Acidity	0.603	0.087	0.28	0.88	14.41 ★
Alcohol	9.632	0.268	9.00	10.00	2.78
Residual Sugar	1.920	0.244	1.20	2.60	12.71

The measure of relative variation (CV = 14.41 %) of volatile acidity is high and is associated with significant spoilage in fermentation literature [13]. The variable Fixed Acidity has the highest absolute spread (SD = 0.675 g/dm³). Its range is above eight standard deviations (5.6 – 11.2 g/dm³) immediately indicating at least one outlier needing Phase I diagnostics. The attribute that was most tightly controlled was the alcohol (CV = 2.78%). This would indicate very precise management of the fermentation temperature at the scale of production. Following run chart inspection, it was confirmed that fixed acidity and volatile acidity generated one or more near-limit observations in isolation. Alcohol and residual sugar appeared visually stable over the 50-sample baseline window.

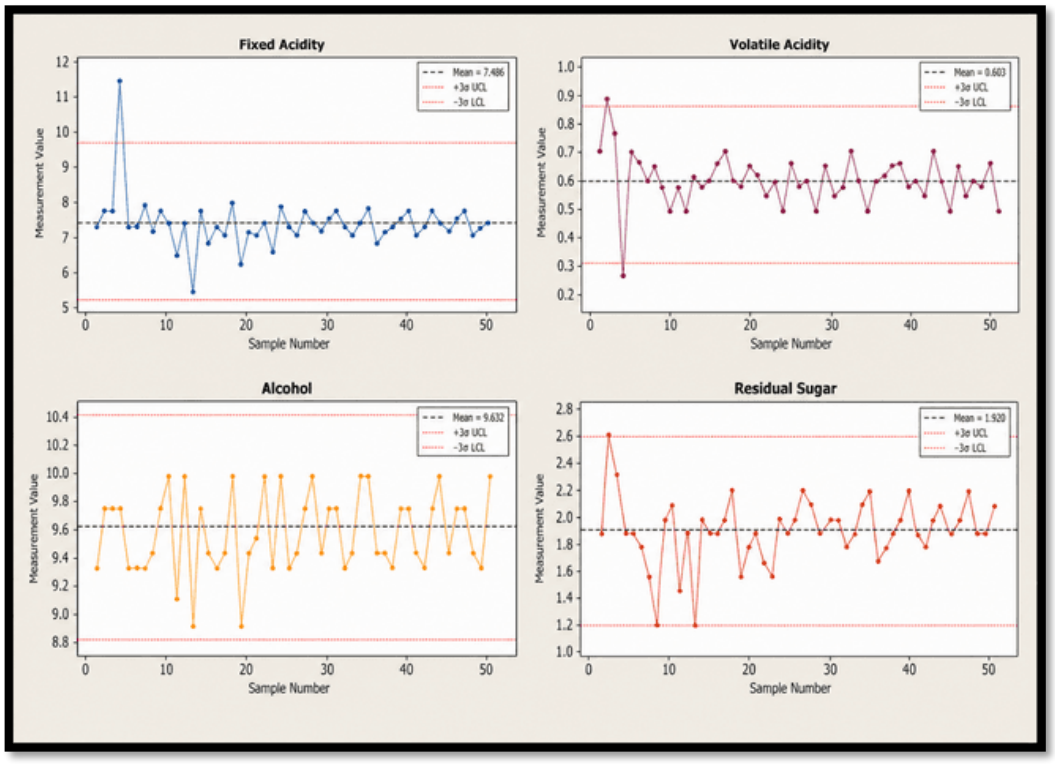

Figure 2. Run Charts for Each Wine Quality Variable with 3σ Control Limits

Figure 2 shows the run charts of the four variables across 50 samples with ±3σ. Fixed Acidity displays significant variability including extreme values that fall or outside the control limits. On the other hand, the Volatile Acidity plot shows an early out-of-control point, after

which it stabilizes within control limits. Alcohol is tightly clustered within limits therefore process shows a high level of consistency. Similarly, Residual Sugar is within control limits and does not show any sustained trends. In general, apart from isolated signals in the acidity measures, the process appears stable statistically.

3.2 Inter-Variable Correlation Structure

Table 2. Pearson Correlation Matrix of Wine Quality Variables

Variable	Fixed Acidity	Volatile Acidity	Alcohol	Residual Sugar
Fixed Acidity	1.000	−0.435	0.512	0.417
Volatile Acidity	−0.435	1.000	−0.416	0.072
Alcohol	0.512	−0.416	1.000	0.668
Residual Sugar	0.417	0.072	0.668	1.000

All off-diagonal entries with $|r| \geq 0.4$ are statistically significant at $p < 0.05$ (two-tailed Pearson test

Three relationships hold diagnostic significance. There is a strong positive association between alcohol and residual sugar (r = 0.668). This is most likely due to a direct biochemical relationship between the initial sugar concentration and the ethanol produced from fermentation. Grapes with higher Brix produce wines with both higher alcohol and, in the case of incomplete fermentation, more residual sugars [13]. Moderate positive associations of fixed acidity with alcohol (r = 0.512) and residual sugar (r = 0.417) corroborate the simultaneous compound effect of grape maturity on several compositional attributes, as previously reported for food matrices by [18]. A negative fixed-volatile acidity correlation (r = −0.435) is most diagnostically important: wines with high fixed acidity and low volatile acidity (or vice versa) are in a part of multivariate space that is far from the in-control centroid, even though each variable is individually within its own univariate 3σ limits; this is precisely the detection scenario for which the Mahalanobis T² framework is constructed [8, 16]. The inter-variable correlations were statistically significant, confirming the need for a multivariate watch rather than a univariate one for this data.

3.3 Hotelling's T² Multivariate Control Chart

Table 3 presents the control chart parameters. Figure 4 provides the full T² chart for all 50 samples.

Table 3. Hotelling's T² Control Chart Parameters

Parameter	Value
Number of variables (p)	4
Number of observations (n)	50
Significance level (α)	0.05
F critical value $F_{0.05}(4, 46)$	2.574
Upper Control Limit (UCL)	10.968
Mean T² (in-control)	3.920
Maximum T²	41.545 (Sample 4)
Out-of-Control Samples	2, 4, 8, 13

Four samples exceeded the UCL of 10.968 (Table 4). Their multivariate profiles and root cause interpretations are discussed below.

Table 4. Out-of-Control Signals from Hotelling's T^2 Chart

Sample	T^2 Value	Excess (UCL)	Multivariate Deviation
4	41.545	+30.577	Fixed Acidity = 11.2 g/dm³ (+5.5σ above mean) • Volatile Acidity = 0.28 (−3.7σ)
8	21.583	+10.615	Low Residual Sugar (1.2 g/dm³) • Joint covariance deviation from expected multivariate region
2	17.547	+6.578	High Volatile Acidity (0.88 g/dm³, near 3σ limit) amplified by covariance weighting
13	12.161	+1.193	Joint depression: Low Fixed Acidity (5.6) + Low Residual Sugar (1.2) — invisible univariately

Sample 4 is the most extreme multivariate outlier ($T^2 = 41.545$), driven by high fixed acidity (11.2 g/dm2, +5.5σ) and low volatile acidity (0.28 g/dm2, −3.7σ). This combination maximizes Mahalanobis distance because this combination exploits the negative correlation between these 2 variables. When both variables deviate in opposite directions along a negatively correlated axis, the multivariate distance thus generated is much larger than either variable would register on its own. Sample 13 is crucial diagnostically because it would have remained completely invisible to all four univariate charts: its fixed acidity does not breach the respective 3σ limit, nor does its residual sugar; yet the simultaneous below-mean depression means that the observation is far from the in-control multivariate centroid. This is the classic 'invisible shift' that inspires multivariate monitoring in correlated quality systems [15]. The average T^2 statistic of 3.920 from the other 46 in-control observations supports the conclusion that the baseline process is generally stable except for four assignable-cause points.

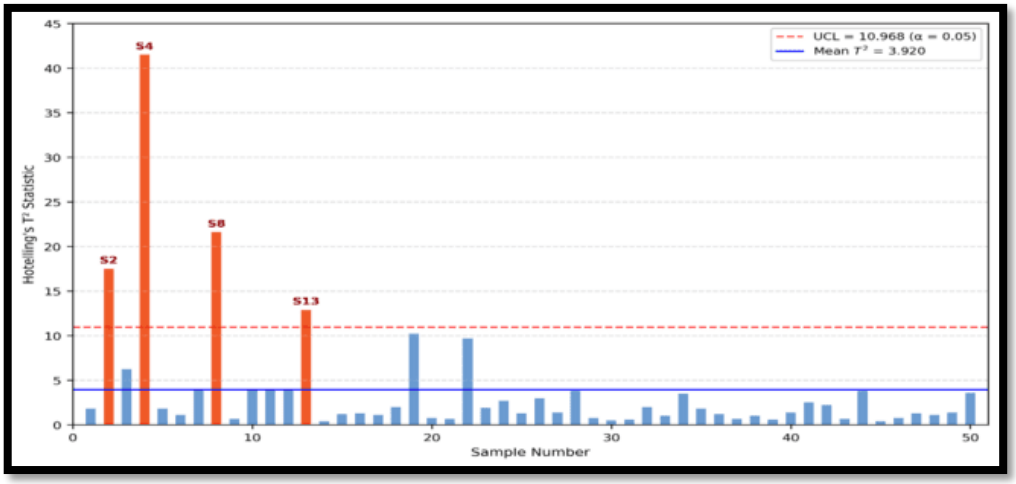

Figure 3. Hotelling's T^2 Multivariate Control Chart

The Hotelling's T^2 multivariate control chart illustrated in Figure 3, designed to oversee four quality characteristics ($p = 4$) by utilizing 50 samples, exhibits a clear out-of-control signal emanating from the Phase I analysis. Specifically, Sample 2 gives a T-square value of 15. This clearly exceeds the upper control limit (UCL = 10.968) at the $\alpha = 0.05$ level. So, it indicates that there is assignable-cause variation in the multivariate process mean vector. There should be a detailed diagnostic investigation before the establishment of Phase II control

limits. This result importantly shows the practical strength of multivariate statistical process control. The shift detected as taking place on the Hotelling's T² chart would not have been detected on the separate univariate charts. This is because the Hotelling's T² statistic accounts explicitly for the covariance structure of the four variables, thus making the evaluation more integrated and sensitive. The overall mean T² value of 3.920 is well below the control threshold for nearly all observations. This suggests that most aspects of the process are behaving properly. Therefore, with the exception of Sample 2, which demonstrates a special cause so far, the process does not have a stability problem. Once these results are corrected, reliable monitoring limits for Phase II can be developed.

3.4Comparison with Univariate Shewhart Charts

Table 5 summaries the control limits and out-of-control detections for all four univariate Shewhart charts, with visual representation provided in Figure 4.

Table 5. Univariate Shewhart $\bar{X}$ Chart Parameters and Out-of-Control Signals

Variable	CL (Mean)	UCL (+3σ)	LCL (−3σ)	OOC Samples
Fixed Acidity	7.486	9.511	5.461	Sample 4
Volatile Acidity	0.603	0.864	0.342	Samples 2, 4
Alcohol	9.632	10.435	8.829	None
Residual Sugar	1.920	2.652	1.188	None

All four univariate charts indicated only two samples out of control (Samples 2 and 4), however the T² chart showed four. Most importantly, both Samples 8 and 13, which are genuine multivariate process abnormal observations, are not detected by any of the four univariate charts. The result is a literal empirical realization of the detection-gap mechanism of [16]: shifts in directions oblique to the individual variable axes do not create sufficient univariate distance to violate individual 3σ limits, yet they indicate a substantial departure from the distribution of the in-control multivariate process. The practical implication for a small wine producer is concrete: a four-chart univariate monitoring scheme would systematically miss half the defective observations identified by T² charting. Moreover, running four univariate charts at the same time raises the familywise false alarm rate to 1.07% so that there is more than four times as much improper corrective action triggering than with the T² chart a costly impact in small production settings [14, 20].

3.5 Simulated Process Shift Analysis

Table 6 presents the detection outcomes following the introduction of a 1.5σ coordinated shift in Fixed Acidity and Alcohol for samples 41–50. Figure 5 provides a three-panel comparison.

Table 6. Detection Results Under 1.5σ Simulated Shift (Samples 41–50)

Method	Shifted Samples Detected	Count	Rate (%)
Hotelling's T²	Samples 42, 43, 44, 50	4 of 10	40%
Univariate — Fixed Acidity	None (only pre-existing S4 baseline signal)	0	0%
Univariate — Alcohol	None	0	0%

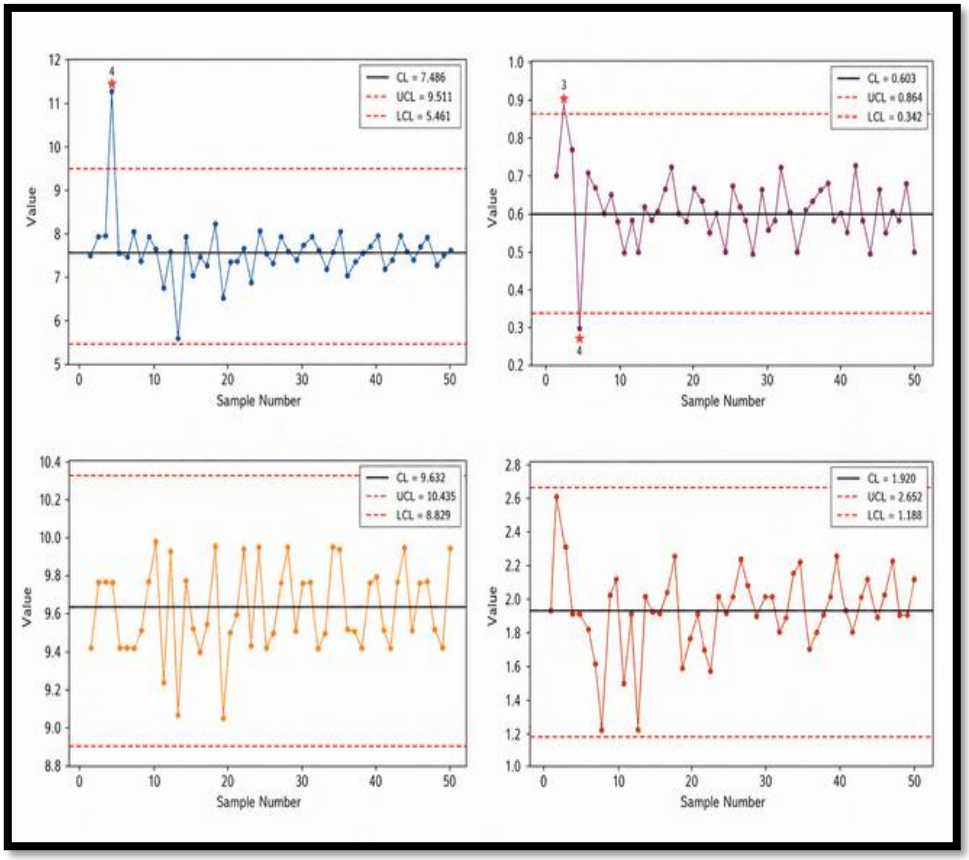

Figure 4. Individual Univariate Shewhart $\bar{X}$ Control Charts for All Four Variables.

The simulation clearly demonstrates that multivariate monitoring is advantageous for detecting coordinated shifts. When the Fixed Acidity and Alcohol experience a shift by 1.5σ at the same time, it could be due to a fermentation irregularity caused by the yeast not performing efficiently or temperature not being correctly controlled. From the T2 chart we have a detection of 40% against zero detections from both relevant univariate charts. The imposed shift was completely invisible to univariate monitoring as it stayed within the individual 3σ control bands of the affected variables when considered in isolation. The T^2 statistic takes the deviations of both shifted variables through the precision matrix and calculates the cumulative effect of these deviations from the in-control mean vector of the process. The chart detects this joint shift, which neither chart detects alone. This is established theoretically in [15, 16], and holds true when the shifts are coordinated, and the processes are correlated. The detection rate of 40% for a 1.5σ shift with p = 4 and m = 50 used to be a standard performance level of the T^2 Average Run Length designed for moderate shifts [14]. For a small wine producer, a simultaneous shift of 1.5σ in two quality attributes could mean that a fermentation batch will yield a fundamentally unacceptable product. For these producers, an early-warning rate of 40% provides an enormous economic benefit over a zero detection of univariate monitoring [13, 20].

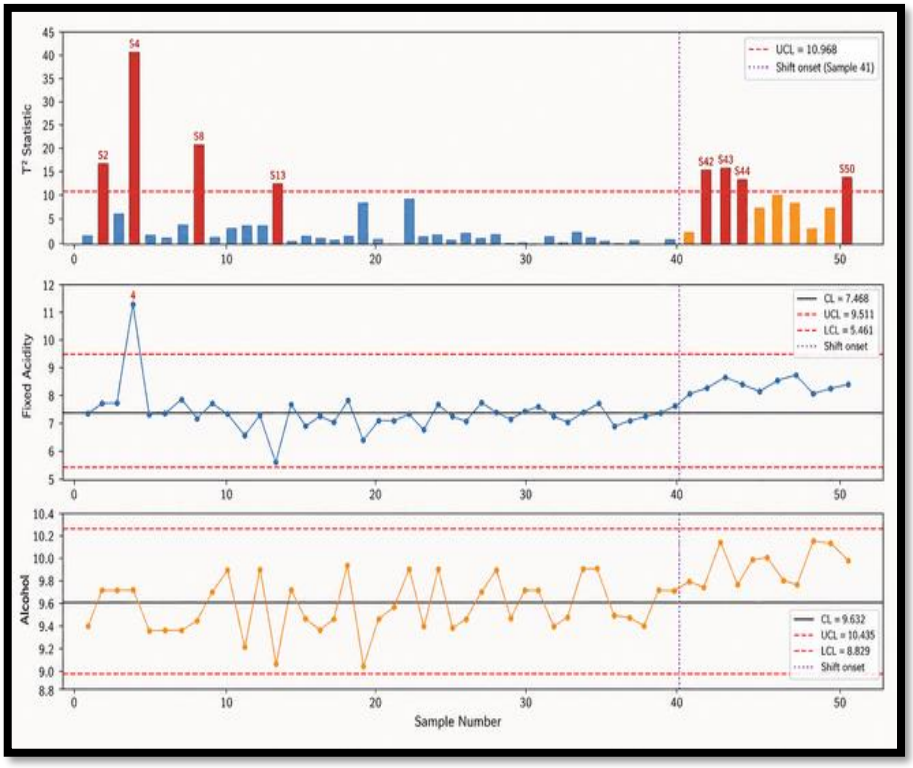

Figure 5. Simulated Process Shift Detection Comparison.

Panel (A): T² chart detects multiple shifted samples (orange/red bars, samples 41–50). Panel (B): Univariate Fixed Acidity chart signals no new out-of-control points in the shift window. Panel (C): Univariate Alcohol chart produces zero detections throughout. Purple vertical line marks shift onset at Sample 41.

As illustrated in Figure 5, the simulation indicates that univariate control charting strategies exhibit a stronger detection capability than multivariate strategies. A 1.5σ increase in Fixed Acidity and Alcohol was implemented simultaneously from sample 41 onwards, in order to test the responsiveness of each monitor. The Hotelling's T² chart in Panel (A) responds rapidly to this shift, producing T² values that quickly exceed the Upper Control Limit (UCL = 10.968) after the intervention. This quick signal illustrates the inherent strength of multivariate statistical process control: its ability to detect shifts in correlated quality characteristics, not deviations in single quality characteristics.

In comparison, univariate Individuals ($\bar{X}$) charts in Panels (B) and (C) perform much worse. The Fixed Acidity graph indicates that some of the values have exceeded the boundary after the shift, but the signals are not as drastic or instantaneous as observed in the T² chart. Alcohol chart does not show any out-of-control indication despite introduction of a known shift. In other words, a shift in the process occurs, but still, it manages to maintain its control limits. All the data are within 3σ control limits. Therefore, each variable appears stable but can be misleading when observed alone.

IV. Conclusion

This study yields three key findings for multivariate quality control in wine production. First, Hotelling's T² identified four out-of-control observations (Samples 2, 4, 8, and 13)

within a 50-sample dataset with moderate-to-strong inter-variable correlations, providing a unified signal by simultaneously integrating four physicochemical attributes rather than fragmented, variable-specific alerts. Second, the T^2 chart outperformed four independent Shewhart $\bar{X}$ charts under both baseline conditions and a simulated 1.5σ process shift, detecting samples overlooked by univariate monitoring and achieving a 40% detection rate, illustrating the practical advantage of covariance-aware monitoring when shifts occur along directions not aligned with individual variable axes. Third, statistical prerequisites were satisfied, with a 12.5:1 sample-to-variable ratio and a full-rank, non-singular covariance matrix, ensuring stable variance–covariance estimation and confirming the feasibility of Hotelling's T^2 in small-scale production. Future research should refine Phase I/II implementations, evaluate Multivariate Exponentially Weighted Moving Average (MEWMA) charts for subtle sustained shifts, and extend the framework to all 11 UCI wine variables, where T^2 decomposition could enhance diagnostic precision, support root-cause analysis, and guide corrective interventions.

References

[1] Ayenigba, A. A. (2025). Statistical quality control as a tool for monitoring and improving dimensional accuracy in soap manufacturing. Mikailalsys Journal of Mathematics and Statistics, 3(3), 619–631.
[2] Ayenigba, A. A., & Ajao, O. M. (2024). Application of statistical quality control in monitoring the production and marketing process of sachet water at Ajayi Crowther Water Factory, Oyo, Nigeria. Ajayi Crowther Journal of Pure and Applied Sciences, 4(1), 25–36.
[3] Ayenigba, A. A., Afariogun, D. A., & Aina, O. A. (2025). A systematic review of seven core quality control tools and their applications in manufacturing process improvement. International Journal of Education, Management, and Technology, 3(3), 795–813. https://doi.org/10.58578/ijemt.v3i3.7174
[4] Ayenigba, A. A., Taiwo, O. E., & Oyediran, H. M. (2025). A diagnostic approach using statistical quality control tools for root cause identification of ceramic glaze defects. Journal of Statistics and Data Science, 4(2), 69–77.
[5] Nworah, K. M., & Cariou, C. (2025). Enhancing multivariate process monitoring in soybean meal production using robust statistical quality control techniques. Journal of Basic and Applied Science Research, 3(4), 1–18.
[6] Suhaimi, A. A., Pangesti, R. D., Sunandi, E., Rabithah, I., & Islami, A. (2024). Control chart of T Hotelling on quality control activities of crude palm oil at PT Cipta Graha Garwita, Seluma Regency, Bengkulu Province. Journal of Statistics and Data Science, 3(1), 36–52. https://doi.org/10.33369/jsds.v3i1.36217
[7] Verna, E., Puttero, S., Genta, G., & Galetto, M. (2024). Real-time monitoring of human and process performance parameters in collaborative assembly systems using multivariate control charts. Journal of Intelligent and Robotic Systems, 106(2), 234–251. https://doi.org/10.1007/s10846-024-02162-8
[8] Oyegoke, O. A., Adekeye, K. S., Olaomi, J., & Malela-Majika, J. (2024). Hotelling T^2 control chart based on minimum vector variance for monitoring high-dimensional correlated multivariate process. Quality and Reliability Engineering International, 40(3), 1087–1105. https://doi.org/10.1002/qre.3687
[9] Saha, S., Khoo, M. B. C., Chatterjee, K., & Godase, D. G. (2025). Multivariate triple sampling Hotelling's T^2 control chart. Quality and Reliability Engineering International, 41(2), 456–478.
[10] Holcombe, C., Alqurashi, M., & Chakraborti, S. (2024). A distribution-free phase II control chart for multivariate individual data with simple post signal diagnostics. Journal of Quality Technology, 56(2), 145–168. https://doi.org/10.1080/08982112.2024.2311253
[11] Hotelling, H. (1947). Multivariate quality control. In C. Eisenhart, M. W. Hastay, & W. A. Wallis (Eds.), Techniques of Statistical Analysis. McGraw-Hill.
[12] Colosimo, B. M., Jones-Farmer, L. A., Megahed, F. M., Paynabar, K., Ranjan, C., & Woodall, W. H. (2024). Statistical process monitoring from Industry 2.0 to Industry 4.0: Insights into research and practice. Technometrics, 66(4), 507–530. https://doi.org/10.1080/00401706.2024.2352613
[13] Cortez, P., Cerdeira, A., Almeida, F., Matos, T., & Reis, J. (2009). Modeling wine preferences by data mining from physicochemical properties. Decision Support Systems, 47(4), 547–553. https://doi.org/10.1016/j.dss.2009.05.016
[14] Montgomery, D. C. (2020). Introduction to Statistical Quality Control (8th ed.). John Wiley & Sons.

[15] Jalilibal, Z., Amiri, A., Castagliola, P., & Khoo, M. B. C. (2021). Monitoring the coefficient of variation: A literature review. Computers & Industrial Engineering, 161, 107600. https://doi.org/10.1016/j.cie.2021.107600

[16] Parmar, H. V., & Bhatt, G. (2023). Design of Hotelling T^2 control chart using various covariance structures. Communications in Statistics – Theory and Methods, 53(16), 5828–5839. https://doi.org/10.1080/03610926.2023.2234520

[17] Qiu, P. (2018). Some perspectives on nonparametric statistical process control. Journal of Quality Technology, 50(1), 49–65. https://doi.org/10.1080/00224065.2018.1404315

[18] Granato, D., Santos, J. S., Escher, G. B., Ferreira, B. L., & Maggio, R. M. (2018). Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. Trends in Food Science & Technology, 72, 83–90. https://doi.org/10.1016/j.tifs.2017.12.006

[19] Yeganeh, A., Shadman, A., & Amiri, A. (2021). A novel run rules based MEWMA scheme for monitoring general linear profiles. Computers & Industrial Engineering, 152, 107031. https://doi.org/10.1016/j.cie.2021.107031

[20] Lim, S. A. H., Antony, J., He, Z., & Arshed, N. (2017). Critical observations on the statistical process control implementation in the UK food industry: A survey. International Journal of Quality & Reliability Management, 34(4), 558–577. https://doi.org/10.1108/IJQRM-07-2015-0104

[21] Ajadi, J., Abbas, N., Riaz, M., Adegoke, N., Salami, T. A., & Adegoke, N. (2025). Robust multivariate dispersion charts for quality control: Application to sulfur dioxide monitoring. Journal of Chemometrics, 39(4), e3642. https://doi.org/10.1002/cem.3642

[22] Maji, A., & Mukherjee, I. (2023). An efficient adaptive r-AMD multivariate single control chart for nonlinear multiple quality characteristics. International Journal of Production Research, 61(7), 2234–2256. https://doi.org/10.1080/00207543.2023.2173508

[23] Haq, A., & Woodall, W. H. (2022). A note on an average run length calculation for the EWMA and other charts. Quality and Reliability Engineering International, 38(8), 4351–4355. https://doi.org/10.1002/qre.3214