

# Examining the Relationship Between Simple Technology Implementation, Data Literacy, and Spectral Preprocessing Accuracy: A SEM-PLS Approach Using the Savitzky–Golay Method in Chemometric Analysis

**Taufik Hudha Nursyafaah<sup>1</sup>, Hulwatul Adzro<sup>1</sup>, Fitri Aditri<sup>1</sup>, Regina Aulia Ramadhani<sup>1</sup>, Neni Alyani<sup>1,2</sup>, M Miftahul Madya<sup>1</sup>**

<sup>1</sup>Lembaga Riset AI Creation (LRAC), Depok, Indonesia

<sup>2</sup>Institut Pemerintahan Dalam Negeri, Sumedang, Indonesia

\*Email corresponding author: [mmiftahulm29@gmail.com](mailto:mmiftahulm29@gmail.com)

Received 7 March 2026

Accepted 1 April 2026

Available Online 30 April 2026

**ABSTRACT:** The advancement of digital technology has significantly influenced chemometric analysis, particularly in spectral data preprocessing. The Savitzky–Golay method is widely used to improve signal quality, while tools such as Orange Data Mining offer user-friendly analytical capabilities for beginner researchers. This study aims to analyze the effect of simple technology implementation on the accuracy of spectral preprocessing and to examine the moderating role of community data literacy. A quantitative approach was employed using SEM-PLS with data collected from 113 respondents through a Likert scale questionnaire. The results show that spectral preprocessing accuracy meets validity and reliability criteria, while simple technology implementation and data literacy constructs do not fully satisfy these requirements. Despite the absence of multicollinearity issues, measurement limitations affect the structural model evaluation. In conclusion, although preprocessing performance is consistent, improvements in measurement indicators are necessary to better assess the relationship between technology implementation and analytical outcomes.

**ABSTRAK:** Perkembangan teknologi digital telah memberikan dampak signifikan terhadap analisis kemometrik, khususnya dalam preprocessing data spektra. Metode Savitzky–Golay banyak digunakan untuk meningkatkan kualitas sinyal, sementara perangkat seperti Orange Data Mining menawarkan kemudahan analisis bagi peneliti pemula. Penelitian ini bertujuan untuk menganalisis pengaruh implementasi teknologi sederhana terhadap akurasi preprocessing spektra serta menguji peran moderasi literasi data komunitas. Penelitian menggunakan pendekatan kuantitatif dengan metode SEM-PLS terhadap 113 responden melalui kuesioner skala Likert. Hasil penelitian menunjukkan bahwa akurasi preprocessing spektra memenuhi kriteria validitas dan reliabilitas, sedangkan konstruk implementasi teknologi sederhana dan literasi data belum sepenuhnya memenuhi kriteria tersebut. Meskipun tidak ditemukan masalah multikolinearitas, keterbatasan pengukuran memengaruhi evaluasi model struktural. Kesimpulannya, meskipun kinerja preprocessing cukup konsisten, diperlukan perbaikan indikator untuk mengkaji hubungan antara teknologi dan hasil analisis secara lebih akurat.

**Keywords:** Chemometrics; Data Literacy; Orange Data Mining; SEM-PLS; Spectral Preprocessing

## 1. INTRODUCTION

The development of digital technology has driven transformation across various fields of data analysis, including chemometrics, which utilizes statistical and mathematical techniques to analyze chemical data. One crucial aspect of chemometric

analysis is spectral data preprocessing, which aims to improve signal quality before further analysis. The *Savitzky–Golay smoothing* method is one of the most widely used preprocessing techniques due to its ability to reduce noise without removing the essential characteristics of the spectrum (Savitzky & Golay, 1964). On the other hand, the emergence of visual-based software such as Orange Data Mining provides ease of use

for novice users in performing data analysis without requiring complex programming skills (Demšar *et al.*, 2013).

The implementation of simple technology such as Orange Data Mining is particularly relevant in the context of beginner research communities that require analytical tools that are easy to use yet still possess adequate analytical capabilities. Ease of installation, a visual *drag-and-drop* interface, and intuitive result interpretation make Orange a promising alternative for chemometric analysis. However, the effectiveness of implementing such simple technology in improving data preprocessing quality, particularly in terms of spectral smoothing accuracy, still requires further empirical investigation (Wold *et al.*, 2001).

Previous studies have extensively examined the adoption and application of data analysis technologies using the Structural Equation Modeling–Partial Least Squares (SEM-PLS) approach, particularly in analyzing relationships among latent variables such as perceived ease of use, perceived usefulness, and analytical outcomes (Venkatesh & Davis, 2000; Hair *et al.*, 2017). Beyond foundational models, recent empirical studies demonstrate that SEM-PLS is widely applied across various domains, including digital technology adoption in education, where factors such as communication, ease of use, effectiveness, and sustainability significantly influence continued technology usage (Pratama *et al.*, 2025). In the field of chemometrics and data-driven education, SEM-PLS has also been used to analyze how perception and knowledge contribute to the development of analytical competence, highlighting its relevance in evaluating data literacy and analytical skills (Alyani *et al.*, 2025a). Furthermore, SEM-PLS has been effectively implemented in broader social and organizational contexts, such as examining the formation of tolerance through leadership, cultural intelligence, and social adaptation (Alyani *et al.*, 2025b), as well as assessing work readiness influenced by

program quality, mentor support, and technology utilization (Adzro *et al.*, 2026). In digital behavior studies, SEM-PLS has been used to evaluate the influence of promotion and social factors on online purchasing decisions through mediating variables such as convenience (Dipura *et al.*, 2026), and to analyze how service awareness and management affect user engagement in literacy-related activities (Alyani *et al.*, 2026a). Additionally, studies have shown its capability in modeling relationships involving work environment, motivation, and competence in shaping workforce readiness (Nursyafaah *et al.*, 2026a). Supporting its application in the field of analytical chemistry, recent research highlights that SEM-PLS can effectively analyze the influence of cognitive understanding on the implementation of method validation, demonstrating that conceptual knowledge—such as understanding validation objectives, parameters (e.g., accuracy and precision), and chemometric principles—plays a significant role in ensuring systematic and reliable analytical practices (Nursyafaah *et al.*, 2026b). In the context of emerging technologies, SEM-PLS has also been widely used to analyze the impact of technology readiness on generative AI adoption and its implications for organizational performance (Nursyafaah *et al.*, 2026c), as well as the role of digital capability in enhancing innovation and operational efficiency (Alyani *et al.*, 2026b), and the influence of perceived ease of use on AI adoption in strengthening competitive advantage (Alyani *et al.*, 2026c). These findings indicate that SEM-PLS is a robust and flexible analytical approach for evaluating complex relationships between technology implementation and performance outcomes. Therefore, in this study, SEM-PLS is employed to evaluate the effect of simple technology implementation on spectral preprocessing accuracy, while also considering the moderating role of community data literacy.

Despite the extensive application of SEM-PLS in analyzing technology adoption,

user behavior, and performance outcomes across various domains, a significant research gap remains in integrating simple technology implementation with technical data processing accuracy within a chemometric context. Most prior studies tend to emphasize behavioral, organizational, or adoption-related factors without explicitly linking them to the quality of analytical outputs, particularly in spectral data preprocessing. In addition, although the *Savitzky–Golay* method is widely recognized for its effectiveness in smoothing spectral data, limited studies have empirically examined how the ease of use and accessibility of analytical tools, such as visual-based platforms like Orange Data Mining can influence the accuracy of preprocessing results. Furthermore, the moderating role of community data literacy, which may determine users' ability to effectively utilize such tools and interpret results, has not been sufficiently explored in relation to preprocessing performance. Therefore, this study aims to fill these gaps by analyzing the effect of simple technology implementation using Orange Data Mining on the accuracy of spectral preprocessing based on the Savitzky–Golay method, while also examining the moderating role of community data literacy using the SEM-PLS approach.

## 2. METHODOLOGY

This study employs a quantitative research approach to examine the relationships among variables related to simple technology implementation and spectral preprocessing accuracy. The research design is explanatory, aiming to test the influence of the independent variable on the dependent variable, as well as the moderating effect of data literacy. Data were collected through a structured questionnaire using a five-point Likert scale ranging from strongly disagree (1) to strongly agree (5). The respondents of this study consisted of 113 individuals from a community of beginner researchers who have experience or exposure to data analysis tools, particularly Orange Data Mining. The variables used in this study

include simple technology implementation as the independent variable, spectral preprocessing accuracy using the Savitzky–Golay method as the dependent variable, and community data literacy as the moderating variable. The indicators of each variable are presented in a separate table.

**Table 1** Variable, Definition, and Indicator

Variable	Type	Code	Indicator
X	Independent	X1	I can install Orange Data Mining easily
		X2	The visual interface of Orange is easy for new users to understand
		X3	The process of inputting analytical data into Orange is easy to perform.
		X4	I can create an analysis workflow using the drag-and-drop system easily.
		X5	The chemometric analysis features in Orange are easy to use.
		X6	The analysis results displayed by Orange are easy to understand.
Y	Dependent	Y1	The smoothing method helps to refine the spectrum.
		Y2	Noise in the spectrum is reduced after preprocessing.
		Y3	Spectral peaks become clearer after preprocessing.
		Y4	The results of spectral data preprocessing are sufficiently stable.
Z	Moderator	Z1	I Understand the basic concepts of data analysis.
		Z2	I am able to read graphs and data visualizations effectively.
		Z3	I am capable of interpreting data analysis results.
		Z4	I have experience using data analysis software.

The data analysis technique used in this study is *Structural Equation Modeling–Partial Least Squares* (SEM-PLS), which is considered suitable for exploratory research and complex models involving latent variables. SEM-PLS is widely applied in studies with relatively small to medium sample sizes and does not require strict assumptions of data normality (Hair et al., 2017). Furthermore, SEM-PLS enables simultaneous analysis of measurement and structural models, making it appropriate for examining both direct and moderating effects among variables in a single framework (Chin, 1998).

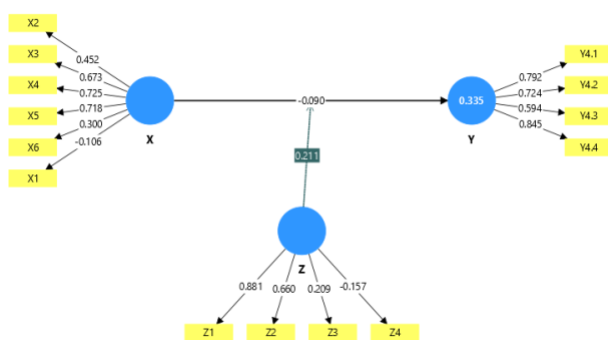
The evaluation of the measurement model (outer model) is conducted to assess the validity and reliability of the constructs. Convergent validity is evaluated using outer loading values, where a threshold of 0.70 is considered acceptable. Multicollinearity among indicators is assessed using the Variance Inflation Factor (VIF), with recommended values below 5. Additionally, the Average Variance Extracted (AVE) is used to confirm construct validity, with a minimum threshold of 0.50. Reliability is measured using Cronbach’s Alpha and composite reliability (rho c), where values above 0.70 indicate satisfactory internal consistency (Hair et al., 2017; Ringle et al., 2015).

technology implementation and spectral preprocessing accuracy. The significance of path coefficients is assessed using a bootstrapping procedure, which provides t-statistics and p-values for hypothesis testing. Through this approach, the study aims to provide empirical evidence on the effectiveness of simple technology implementation in improving preprocessing accuracy within a chemometric analysis framework (Chin, 1998; Ringle et al., 2015).

### 3. RESULT

The evaluation of the measurement model (outer model) was conducted to assess the validity and reliability of the constructs using outer loading, Average Variance Extracted (AVE), Cronbach’s Alpha, composite reliability (rho\_c), and Variance Inflation Factor (VIF). The results indicate that the dependent variable, spectral preprocessing accuracy (Y), meets the criteria for both validity and reliability, with an AVE value of 0.555, Cronbach’s Alpha of 0.738, and composite reliability of 0.831. Most indicators under this construct demonstrate acceptable outer loading values above the recommended threshold of 0.70, except for one indicator which is slightly below but still considered acceptable in exploratory research (Henseler et al., 2015).

In contrast, the independent variable, simple technology implementation (X), does not fully meet the required criteria for construct validity and reliability. The AVE value of 0.300 indicates insufficient convergent validity, while Cronbach’s Alpha (0.678) and composite reliability (0.645) fall below the recommended threshold of 0.70. Only two indicators within this construct achieve acceptable loading values, while the remaining indicators show low or even negative loadings, suggesting weak representation of the latent construct. Similarly, the moderating variable, community data literacy (Z), also fails to meet validity and reliability standards, with AVE (0.320), Cronbach’s Alpha



**Fig. 1** Model Diagram and Intervariable Relationship

Furthermore, the structural model (inner model) is analyzed to examine the relationships between variables and to test the research hypotheses. The moderating effect of community data literacy is evaluated to determine whether it strengthens or weakens the relationship between simple

(0.340), and composite reliability (0.483) indicating poor construct measurement.

Despite these limitations, the multicollinearity test shows that all indicators across constructs have VIF values below 5, indicating no significant collinearity issues among indicators. This suggests that although some constructs lack internal consistency, the indicators themselves do not exhibit redundancy or strong intercorrelation. Therefore, the model remains statistically acceptable in terms of multicollinearity but requires refinement in terms of measurement quality (Hair et al., 2011).

Construct & Indicator	Loading Factor	AVE	Cronbach's Alpha	rho <sub>c</sub>	VIF
<b>X</b>		0.300	0.678	0.645	
X1	-0.106				1.119
X2	0.452				1.135
X3	0.673				1.321
X4	0.725				1.611
X5	0.718				1.512
X6	0.300				1.207
<b>Y</b>		0.555	0.738	0.831	
Y1	0.792				1.561
Y2	0.724				1.510
Y3	0.594				1.364
Y4	0.845				1.595
<b>Z</b>		0.320	0.340	0.483	
Z1	0.881				1.108
Z2	0.660				1.138
Z3	0.209				1.058
Z4	-0.157				1.065
Z x X	1.000				1.000

Construct & Indicator	Description
<b>X</b>	<b>Not valid, not reliable</b>
X1	Not valid
X2	Not valid
X3	Not valid
X4	Valid
X5	Valid
X6	Not valid
<b>Y</b>	<b>Valid &amp; reliable</b>
Y1	Valid
Y2	Valid
Y3	Not valid
Y4	Valid
<b>Z</b>	<b>Not valid &amp; not reliable</b>
Z1	Valid
Z2	Not valid

Z3	Not valid
Z4	Not valid
Z x X	Valid

#### 4. DISCUSSION

The findings of this study highlight that spectral preprocessing accuracy using the Savitzky–Golay method can be reliably measured, supporting previous literature that emphasizes the robustness of smoothing techniques in improving signal quality in chemometric analysis. The strong performance of the dependent variable indicates that respondents are able to consistently perceive improvements in spectral quality after preprocessing, particularly in terms of noise reduction and peak clarity. This aligns with the theoretical understanding that preprocessing plays a critical role in enhancing the interpretability of spectral data.

However, the results reveal significant limitations in the measurement of simple technology implementation. The low validity and reliability values suggest that the indicators used may not adequately capture the construct, or that respondents have varied perceptions regarding the usability of Orange Data Mining. This finding contrasts with prior studies that highlight ease of use and perceived usefulness as strong determinants of technology adoption (Davis, 1989). One possible explanation is that while Orange provides a user-friendly interface, differences in user experience, familiarity with data analysis tools, and technical background may influence how respondents evaluate its usability.

Similarly, the moderating variable of community data literacy demonstrates weak measurement properties, indicating that respondents' self-assessment of their data literacy may not be consistent or well-defined. This suggests that data literacy is a complex construct that may require more comprehensive indicators to capture users' actual analytical capabilities. Previous studies have emphasized that data literacy significantly influences the effectiveness of technology utilization and analytical outcomes (Alharthi, 2021), yet this study indicates that its role as a moderator cannot



be fully validated due to measurement limitations.

Overall, the results imply that while the technical outcome (preprocessing accuracy) is well-established, the human and technological factors influencing it require further refinement. This finding supports the research gap identified earlier, where the integration between technology usability and analytical accuracy remains underexplored. Future research should focus on improving the measurement model by refining or eliminating weak indicators, as well as expanding the construct of data literacy to better capture users' competencies. By addressing these limitations, future studies can provide a more comprehensive understanding of how simple technology implementation influences analytical performance in chemometric applications (Davis, 1989).

## 5. CONCLUSION

This study concludes that spectral preprocessing accuracy using the Savitzky–Golay method can be measured reliably and demonstrates consistent improvement in spectral quality. However, the constructs of simple technology implementation and community data literacy do not meet the required validity and reliability criteria, indicating limitations in how these variables are measured. While no multicollinearity issues were found, the results suggest that the effectiveness of simple technology, such as Orange Data Mining, in influencing preprocessing accuracy cannot be fully confirmed due to measurement weaknesses. Therefore, future research is recommended to refine the measurement indicators and improve construct validity to better capture the relationship between technology implementation and analytical performance in chemometric studies.

## REFERENCES

- Adzro, H., et al. (2026). Analysis of interns' work readiness: The roles of program quality, mentor support, and project management tools using SEM-PLS. *J-GADS: Journal of Governance and Applied Data Science*, 1(1).
- Alharthi, S. W. S. (2021). Data literacy and its impact on decision-making effectiveness. *International Journal of Information Management*, 58, Article 102–xxx.
- Alyani, N., et al. (2025a). Analysis of perception and knowledge on chemometric competence among students and practitioners using SEM-PLS. *Jurnal Manajemen Sumber Daya Aparatur*, 13(2).
- Alyani, N., et al. (2025b). Building a culture of tolerance through sustainable leadership, cultural intelligence, and social adaptation: Evidence from a multicultural community in Indonesia. *Khazanah Sosial*, 7(4).
- Alyani, N., et al. (2026a). Service awareness and management effects on library visit frequency mediated by user interest using SEM-PLS. *J-GADS: Journal of Governance and Applied Data Science*, 1(1).
- Alyani, N., et al. (2026b). Pengaruh digital capability terhadap innovation capability dan operational efficiency startup melalui pemanfaatan AI generatif. *J-KBDI: Jurnal Kecerdasan Buatan dan Data Science Industri*, 1(1).
- Alyani, N., et al. (2026c). Pengaruh perceived ease of use terhadap AI generative adoption dalam meningkatkan competitive advantage startup digital. *J-KBDI: Jurnal Kecerdasan Buatan dan Data Science Industri*, 1(1).
- Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). Lawrence Erlbaum Associates.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Demšar, J., et al. (2013). Orange: Data mining toolbox in Python. *Journal of Machine Learning Research*, 14, 2349–2353.

- Dipura, I. F., et al. (2026). The influence of promotion and social influence on online purchase decisions mediated by convenience. *J-GADS: Journal of Governance and Applied Data Science*, 1(1).
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Sage.
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, 19(2), 139–152.
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135.
- Nursyafaah, T. H., et al. (2026a). The effect of work environment and motivation on work readiness through competence using a SEM-PLS approach. *J-GADS: Journal of Governance and Applied Data Science*, 1(1).
- Nursyafaah, T. H., et al. (2026b). Pengaruh technology readiness terhadap adopsi AI generatif dan dampaknya terhadap kinerja startup menggunakan SEM-PLS. *J-KBDI: Jurnal Kecerdasan Buatan dan Data Science Industri*, 1(1).
- Nursyafaah, T. H., et al. (2026). Pemanfaatan SEM-PLS dalam menganalisis variabel kognitif terhadap validasi metode pada bidang kimia. *J-TP MIPA: Jurnal Transformasi Pendidikan MIPA*, 1(1).
- Pratama, M. O., et al. (2025). Analisis kebutuhan dan minat dalam pemanfaatan teknologi digital di SMK Depok menggunakan SEM-PLS. *Jurnal Terapan Informatika Nusantara*, 6(6), 635–643.
- Ringle, C. M., Wende, S., & Becker, J.-M. (2015). *SmartPLS 3*. SmartPLS GmbH.
- Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8), 1627–1639.
- Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 186–204.
- Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2), 109–130.