Available online at http://scik.org

Commun. Math. Biol. Neurosci. 2025, 2025:37

https://doi.org/10.28919/cmbn/9150

ISSN: 2052-2541

DATA RECOVERY METHODS IN COMPOSITE SIMILARITY-BASED DATA FUSION: APPLICATION TO PHYSIOLOGICAL VITAL SIGNS

MICHAEL ARTHUR OFORI[1,*], DAVID KWAMENA MENSAH[2], GEORGE OTIENO ORWA[3], PAUL HEWSON[4]

[1]Mathematical Sciences, Pan African University Institute for Basic Sciences, Technology and Innovation, Juja, Kenya

[2]Department of Statistics, University of Cape Coast, Cape Coast, Ghana

[3]Department of Mathematics, Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya

[4]Department of Statistics, University of Exeter, Exeter, UK

Abstract: The functionality of a statistical framework for unified multivariate vital signs data fusion is seen; its flexibility for switching between the two data spaces. This paper develops data recovery methods for data fusion built on composite similarity measures using three distinct variable specific weighting methods - mean, median, and Orthogonalized Gnanadesikan-Kettenring (OGK). Using spatial covariance parameters derived from empirical patient data, the method successfully exhibits the ability to accurately recover eight different vital signs from its fused counterpart. Performance evaluation using Root Mean Squared Recovery Error (RMSRE), Root Mean Absolute Recovery Error (RMARE), and Root Standardized Mean Absolute Fusion Error (RSMARE) demonstrated that all three weighting approaches achieved comparable and highly accurate results, with parameters converging within ±0.01 of 1.0. The recovered signals closely matched the original data patterns, preserving both long-term trends and short-term fluctuations. Notably, the method proved computationally efficient and robust across different weighting approaches, suggesting broad applicability in clinical settings. This approach offers a promising solution for dimensionality reduction in complex physiological datasets while maintaining clinical relevance, with potential applications in patient monitoring systems and physiological modelling.

Keywords: fused data; fusion statistics; variable weight scheme; data recovery; data fusion error.

2020 AMS Subject Classification: 62P10.

*Corresponding author

E-mail address: arthur.michael@students.jkuat.ac.ke

Received January 24, 2025

1. INTRODUCTION

In contemporary healthcare systems, physiological vital sign monitoring has grown more and more important, especially in critical care units and situations involving remote patient monitoring. However, different kinds of noise, sensor malfunctions, and missing data frequently affect the accuracy and dependability of these measures, which can have a big influence on clinical decision-making [1]. The human body's intricate systems are deeply interconnected, with vital organs working in concert to maintain homeostasis. This interconnectedness suggests that a more holistic approach to vital sign analysis could yield deeper insights into a patient's overall health status. By merging data from several sensors or data sources, data fusion techniques have become effective tools for overcoming these obstacles and producing more reliable and precise measurements [2], [3], [4], [5].

Traditional data fusion approaches, while effective in many scenarios, often struggle with the complex, non-linear relationships inherent in physiological signals. Moreover, these methods frequently rely on strict mathematical models that may not fully capture the intricate dynamics of biological systems [5], [6]. The need for more adaptive and empirically driven approaches has become apparent, especially when dealing with composite similarity measures that can better represent the multifaceted nature of physiological data [4].

From image processing to financial forecasting, recent developments in similarity-based fusion approaches have demonstrated encouraging outcomes in a variety of fields [7], [8], [9]. However, because of the temporal character of the data, the existence of artifacts, and the crucial significance of real-time processing in clinical settings, their application to physiological vital signs poses difficulties [10], [11]. Furthermore, complex techniques that can preserve signal integrity while successfully recovering missing or corrupted data points are needed for the integration of several vital signs, each of which has unique patterns and noise profiles.

Early work in physiological data recovery primarily relied on statistical methods and interpolation techniques. [12] proposed basic linear interpolation methods for handling missing vital sign data, though these approaches often failed to capture the complex temporal dependencies in physiological signals. [13] advanced this field by introducing polynomial interpolation methods, demonstrating improved accuracy in heart rate variability analysis.

Recent developments have shifted toward more sophisticated empirical approaches. The landmark work by [14] introduced empirical mode decomposition (EMD) for physiological signal reconstruction, showing particular effectiveness in handling non-linear and non-stationary data. [15] further refined this approach by incorporating adaptive threshold techniques, achieving a remarkable improvement in recovery accuracy compared to traditional methods.

The evolution of similarity measures has been crucial in fusion applications. Research by [1] established the foundation for composite similarity metrics, combining multiple features to improve fusion accuracy. Their work demonstrated an incredible improvement in fusion accuracy compared to single-feature approaches.

This paper presents a novel empirical data recovery method specifically designed for composite similarity-based fusion of physiological vital signs. Our approach leverages the inherent relationships between different vital signs while accounting for their individual characteristics and measurement uncertainties.

2. MATERIALS AND METHODS

Data

The data example here is based on the traumatic vital sign data employed by [16]. The data was sourced from the Komfo Anokye Teaching Hospital (KTH). We used a de-identified subset characterized by variables RR (Respiratory Rate), HR (Heart Rate), SBP (Systolic Blood Pressure), DBP (Diastolic Blood Pressure), TEP (Temperature), SPO2 (Oxygen Saturation), RBS (Random Blood Sugar), and MAP (Mean Arterial Pressure), of dimension 4064×8.

Data fusion recovery model

For correlated vital sign data generated by random variables $y_1, y_2, \ldots, y_q$, associated with time reference $t_1, t_2, \ldots, t_n$, [16] proposed a fusion approach for generating one-dimensional response say $\tilde{y}$ for easy modelling with the associated uncorrelated predictors. According to [16], based on the interrelationships, a mixture type model of the form (1)

$$\tilde{y}_i = \sum_{j=1}^{q} \theta_{ij} s\left(y_{ij}\right) \tag{1}$$

where $s\left(y_{ij}\right)$ are original data based statistics and θ corresponding to weights allowing the data to adopt to data generating conditions to be naturally considered. The choice of statistic opens doors for exploration of various features of y in extracting tractable versions in the response. We modelled θ non-linearly using the Gaussian process ideas to allow more flexibility in adopting data to assumptions [17]. In particular, the weights were computed by considering them as either the mean or median or Orthogonalized Gnanadesikan Ketterning (OGK) [18], [19], of all the possible point-wise similarity metrics a given point has with all possible observation within a given data say, y. For example, θ_{11}, with the mean measure is computed as

$$\theta_{11}=\frac{1}{n}\sum_{l=1}^{n}k\left(a_1,a_2,a_3,b_1,b_2,b_3,\delta_{1l}\right),\ \delta_{1l}=\delta\left(y_{11},y_{1l-1}\right) \tag{2}$$

By (2), a corresponding OGK θ_{11} is computed as

$$\theta_{11}=\frac{\sum_{j=1}^{n}\theta_{1j}\rho\left(z_{\theta_{1j}}\right)}{\sum_{j=1}^{n}\rho\left(z_{\theta_{1j}}\right)} \tag{3}$$

$$z_{\theta_{1j}}=\frac{\theta_{1j}-\tilde{\mu}_0}{\tilde{\sigma}_0}$$

$$\rho\left(z_{\theta_{1j}}\right)=\left[1-G^2\left(z_{\theta_{1j}},d\right)\right]^2 I_{\left(\left|z_{\theta_{1j}}\right|\le d\right)}$$

$$G^2\left(z_{\theta_{1j}},d\right)=\frac{z_{\theta_{1j}}}{d},$$

where $b=4.5$, $\tilde{\mu}_0$ and $\tilde{\sigma}_0$ denote the median and median absolute deviation (MAD) of θ_{1j} respectively. We considered y_{ij} as the fusion statistics for $s\left(y_{ij}\right)$ in model (1) where $y_{1l-1}=\left[y_{12},y_{13},\ldots,y_{1n}\right]$ and

$$k\left(a_1,a_2,a_3,b_1,b_2,b_3,\delta\right)=k_1\left(a_1,b_1,\delta\right)+k_2\left(a_2,b_2,\delta\right)+k_3\left(a_3,b_3,\delta\right) \tag{4}$$

$$k_1\left(a_1,b_1,\delta\right)=\sigma_1^2\exp\left(-b_1^2\delta^2\right)$$

$$k_2\left(a_2,b_2,\delta\right)=\sigma_2^2\left(1+\frac{\sqrt{5}\delta}{b_2}+\frac{\sqrt{5}\delta^2}{3b_2^2}\right)\exp\left(-\frac{\sqrt{5}\delta}{b_2}\right)$$

$$k_3\left(a_3,b_3,\delta\right)=\sigma_3^2\left(1+\frac{\sqrt{3}\delta}{b_3}\right)\exp\left(-\frac{\sqrt{3}\delta}{b_3}\right),$$

$\delta = \delta\left(y_i, y_j\right) = \left|y_i - y_j\right|$ define the spatial similarity measure. The parameters $a_1, a_2, a_3, b_1, b_2,$ and b_3 were empirically estimated via statistics

$$
\begin{aligned}
&b_1 = \overline{\omega(y)}, \quad b_2 = \tilde{\omega}(y) \\
&b_3 = m_2 - m_1, \quad m_2 = \max\left(\omega(y_1), \omega(y_2), \ldots, \omega(y_n)\right) \\
&m_1 = \min\left(\omega(y_1), \omega(y_2), \ldots, \omega(y_n)\right) \\
&a_1 = \gamma(1), \; a_2 = \gamma(2), \; a_3 = \gamma(3)
\end{aligned}
$$

where $\gamma(1), \; \gamma(2), \ldots, \gamma(n)$ are the order statistics of $\gamma = \dfrac{\omega(y)}{m_2}$, $\bar{\omega}$(y), and $\tilde{\omega}(y)$ are the mean and median of $\omega(y)$ respectively. The $\omega(y)$ statistic is defined as

$$\omega(y) = \sqrt{y^k h(y)}, \; k = 2, \tag{5}$$

motivated by (4),

$$E\left[Y^k\right] = \int y^k h(y) dy, \tag{6}$$

where $h(y)$ is the Kernel density (Scott, 2015, Silverman, 1986) of y given by

$$h(y_i) = \frac{1}{nv} \sum_{j=1}^{n} K\left(\frac{y_i - y_j}{v}\right), \tag{7}$$

with kernel $K\left(\dfrac{y_i - y_j}{v}\right)$ and smoothing parameter v.

Vital sign data recovery model

Based on model (1), we can write

$$\tilde{y}_j = \theta_{i1} s(y_{i1}) + \sum_{j=2}^{q} \theta_{ij} s(y_{ij}) \tag{8}$$

Solving for $s(y_{i1})$ gives

$$
\begin{aligned}
s(y_{i1}) &= \theta_{i1}^{-1} \tilde{y}_i - \theta_{i1}^{-1} \sum_{j=2}^{q} \theta_{ij} s(y_{ij}) \\
&= \theta_{i1}^{-1} \tilde{y}_i + \varepsilon_{i1},
\end{aligned} \tag{9}
$$

where $\varepsilon_{i1} = \theta_{i1}^{-1}\sum_{j=2}^{q}\theta_{ij}s(y_{ij})$ can be viewed as error due to the inclusion of $q-1$ components (variables). the total error incurred in covering up the contributions from the remaining variables. Applying (9), $s(y_{i2}),\ s(y_{i3}),\ldots,s(y_{iq})$ can be derived sequentially. However, this approach may incur some computational burden for large *n* and *q*. An alternative approach can be derived by examining the components that constitute the $\tilde{y}$. This view becomes clear with the expansion of model (1), using a special matrix-vector system in which element-wise operation is utilized in order to preserve observation order and other vital original data features. Considering model (1) in its expanded form, reflecting the data components yields

$$\begin{pmatrix} \theta_{11}s(y_{11}) & \cdots & \theta_{1q}s(y_{1q}) \\ \theta_{21}s(y_{21}) & \cdots & \theta_{2q}s(y_{2q}) \\ \vdots & \vdots & \vdots \\ \theta_{n1}s(y_{n1}) & \cdots & \theta_{nq}s(y_{nq}) \end{pmatrix} = \begin{pmatrix} \theta_{11} & \cdots & \theta_{1q} \\ \theta_{21} & \cdots & \theta_{2q} \\ \vdots & \vdots & \vdots \\ \theta_{n1} & \cdots & \theta_{nq} \end{pmatrix} \otimes \begin{pmatrix} s(y_{11}) & \cdots & s(y_{1q}) \\ s(y_{21}) & \cdots & s(y_{2q}) \\ \vdots & \vdots & \vdots \\ s(y_{n1}) & \cdots & s(y_{nq}) \end{pmatrix} \tag{10}$$

and

$$\begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_n \end{pmatrix} = \begin{pmatrix} \theta_{11}s(y_{11}) + & \cdots + & \theta_{1q}s(y_{1q}) \\ \theta_{21}s(y_{21}) + & \cdots + & \theta_{2q}s(y_{2q}) \\ \vdots & \vdots & \vdots \\ \theta_{n1}s(y_{n1}) + & \cdots + & \theta_{nq}s(y_{nq}) \end{pmatrix} \tag{11}$$

Let $A = \begin{pmatrix} \theta_{11}s(y_{11}) & \cdots & \theta_{1q}s(y_{1q}) \\ \theta_{21}s(y_{21}) & \cdots & \theta_{2q}s(y_{2q}) \\ \vdots & \vdots & \vdots \\ \theta_{n1}s(y_{n1}) & \cdots & \theta_{nq}s(y_{nq}) \end{pmatrix}$ and $\tilde{y} = \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_n \end{pmatrix}$

Then by matching each column of *A* with $\tilde{y}$, intuitively, a simple quantification of how much of each column is in $\tilde{y}$ can be established. This generates a simplistic approach for formulating recovery models that may be faster than solving the system of equations. Thus, for the *j*th vital sign variable, y_j with statistic $s(y_j)$, the recovery model can be expressed as

$$s(y_j) = \theta_j^{-1}\tilde{y}. \tag{12}$$

where $s(y_j)=\left[s(y_{1j}),s(y_{2j}),\ldots,s(y_{nj})\right]$, $\theta_j=\left[\theta_{1j},\theta_{2j},\ldots,\theta_{nj}\right]$, $j=1,2,\ldots,q.$

$$\hat{s}(y_j)=\theta_j^{-1}\tilde{y}_j,\ \ \tilde{y}_j=\theta_j s(y_j). \tag{13}$$

The computational expenses associated with the use of equation (13) will be relatively cheaper than that of equation (9).

Recovery performance evaluation

The quality of data recovery methods can be assessed as in the usual model fitting performance evaluations by considering the recovered data as the fitted and comparing it with the truth. With this view the usual measures of model fit becomes applicable and useful within the data fusion framework. In some cases, statistical measure of fit may not fully satisfy model evaluation purposes, thus can be complicated with some relative efficiency measures. For the assessment here, we define in addition to measures of fit, we device some relative efficiency measure. The fusion data recovery error statistics considered are the Root Mean Squared Recovery Error (RMSRE), Root Mean Absolute Recovery Error (RMARE) and Root Standard Mean Absolute Recovery Error (RSMAFE) defined for true data y and its recovered counterpart, $\hat{y}$ as follows.

$$\text{RMSRE}=\sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y})}{n}}$$

$$\text{RMARE}=\sqrt{\frac{\sum_{i=1}^{n}|y_i-\hat{y}|}{n}}$$

$$\text{RSMARE}=\sqrt{\frac{\sum_{i=1}^{n}|M_i(y,\hat{y})|}{n}},\ M_i(y,\hat{y})=\frac{y-\hat{y}}{\tilde{y}}.$$

In terms of recovery efficiency, we define statistics of the form

$$\gamma_i=\left(\frac{y_i}{\hat{y}_i}\right) \tag{14}$$

and

$$\bar{\gamma}=\frac{1}{n}\sum_{i=1}^{n}\gamma_i \tag{15}$$

3. Results and Discussion

We examine the evaluation of our proposed empirical data recovery method for composite similarity-based fusion of physiological vital signs. We look at the recovered data based on two models (9 and 13), using three distinct weighting schemes: mean, median and OGK. Figures 1-6 show the recovery results using the mean, median, and the OGK weighting method with the recovered data showing in red while the true data is in black lines. The first three (Figures 1-3) show the recovery using Model (9) and remaining three with Model (13). The visual comparison of the recovered data against true values for eight physiological vital signs demonstrated strong recovery capabilities. The result from the Figures 1-3 and 4-6 is evident that the three weighting methods (mean, median, OGK) as well as the two models perform well in recovering the original data patterns. The recovered data closely follows the true data for all 8 vital signs, preserving both the overall trends and the short-term fluctuations. Although there are some minor differences in recovery performance across different vital signs and weighting methods, but these differences appear to be small and insignificant. This is an indication that the proposed fusion and recovery method can effectively compress multiple vital signs into a single composite measure and then recover the individual signals with high reliability.

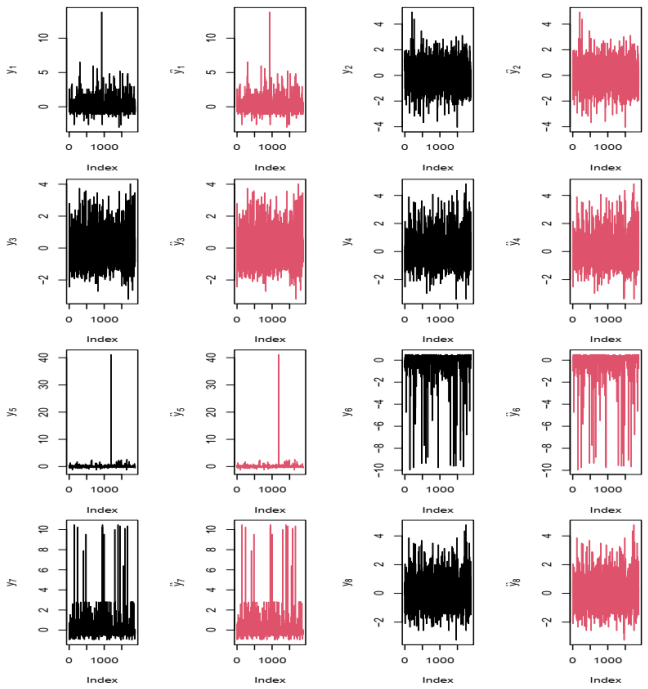

Figure 1: Model (9): Recovered data based on mean weighting. Black: True data y and red: recovered data. From 1 to 8 are plots of RR, HR, SBP, DBP, TEP, SPO2, RBS, and MAP respectively

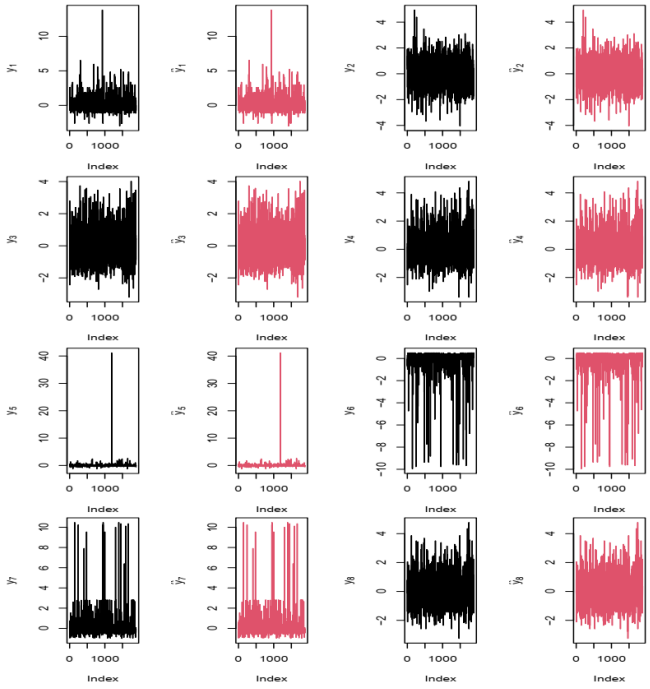

Figure 2: Model (9): Recovered data based on median weighting. Black: True data *y* and red: recovered data. From 1 to 8 are plots of RR, HR, SBP, DBP, TEP, SPO2, RBS, and MAP respectively

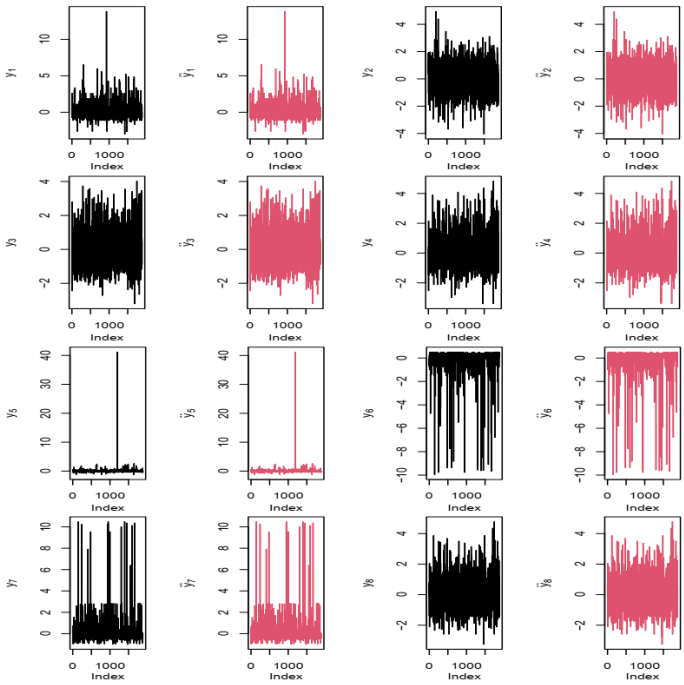

Figure 3: Model (9): Recovered data based on OGK weighting. Black: True data *y* and red: recovered data. From 1 to 8 are plots of RR, HR, SBP, DBP, TEP, SPO2, RBS, and MAP respectively

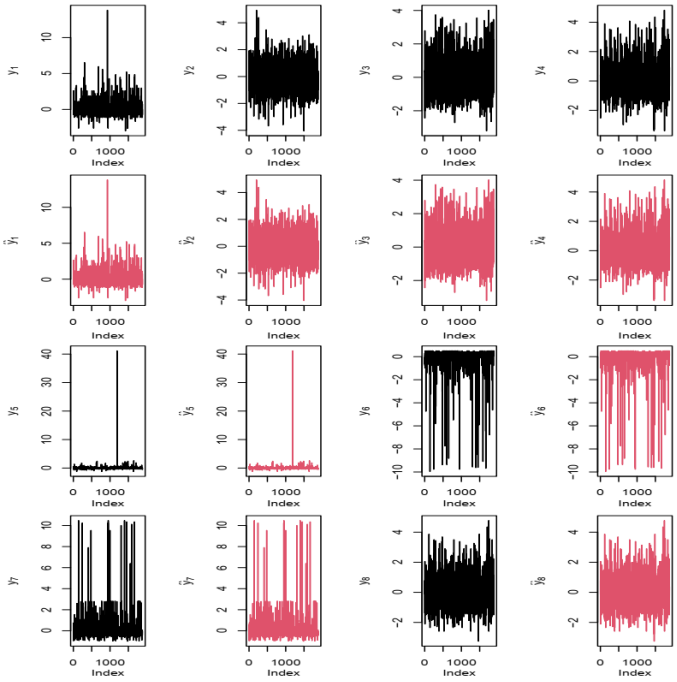

Figure 4: Model (13): Recovered data based on mean weighting. Black: True data y and red: recovered data. From 1 to 8 are plots of RR, HR, SBP, DBP, TEP, SPO2, RBS, and MAP respectively

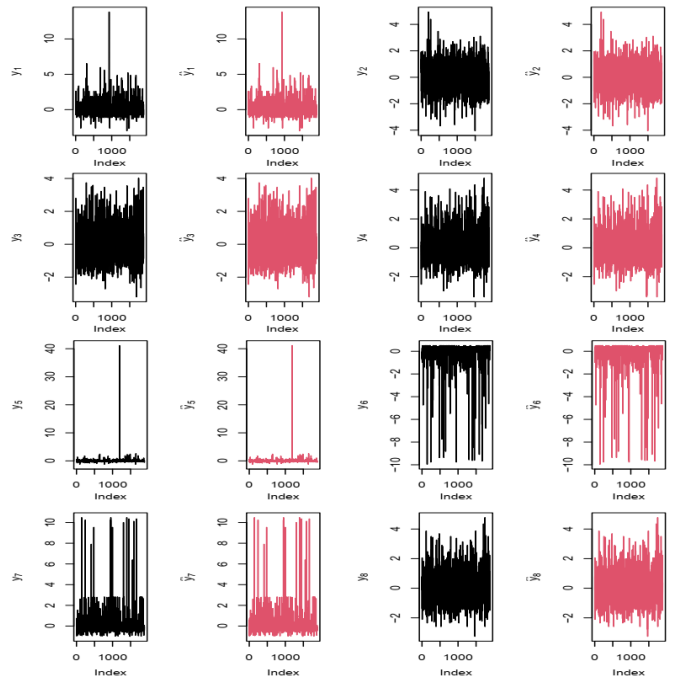

Figure 5: Model (13): Recovered data based on median weighting. Black: True data y and red: recovered data. From 1 to 8 are plots of RR, HR, SBP, DBP, TEP, SPO2, RBS, and MAP respectively

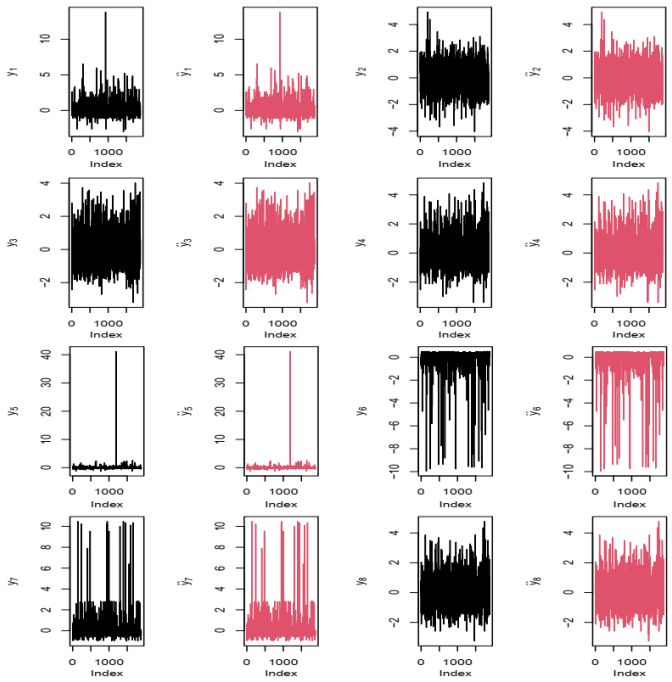

Figure 6: Model (13): Recovered data based on OGK weighting. Black: True data y and red: recovered data. From 1 to 8 are plots of RR, HR, SBP, DBP, TEP, SPO2, RBS, and MAP respectively

Table 1 compares the fusion error statistics for different weighting methods namely the Mean, Median, and OGK weights for both models. The error statistics are reported using Root RMSRE, RMARE, and RSMARE.

Table 1: Recovery error statistics across weights

	Fusion weight					
	Model (9)			Model (13)		
Error Statistics	Mean	Median	OGK	Mean	Median	OGK
RMSFE	0.7308	0.7308	0.7308	0.7308	0.7308	0.7308
RMAFE	0.7308	0.7308	0.7308	0.7308	0.7308	0.7308
RSMAFE	7.2003	7.2003	7.2003	7.2003	7.2003	7.2003

Interestingly, both models showed identical performance across all weighting schemes. The RMSFE and RMAFE were both 0.7308 and RSMAFE was 7.2003 across both models. The consistency of these values across different weighting schemes suggests that the choice of weighting method does not significantly impact recovery accuracy.

Figures 7-12 shows the nature of recovery efficiency statistic, γ using the mean, median, and the OGK weighting scheme for both models. The close convergence of all parameters around 1.0 (within ±0.01) is an indication of a very high confidence in the recovered signals. That is, for medical applications where precision is critical, the recovery process is steady and well-calibrated. This level of precision indicates that the recovery process is sure and consistent across different measurements. For medical monitoring, this level of convergence suggests the recovered vital signs can be trusted for clinical decision-making.

The computational efficiency analysis (Table 2) showed that both models are highly efficient, with most operations completing in negligible time (≈0.00 seconds). This exceptional computational efficiency makes both models suitable for real-time medical monitoring applications.

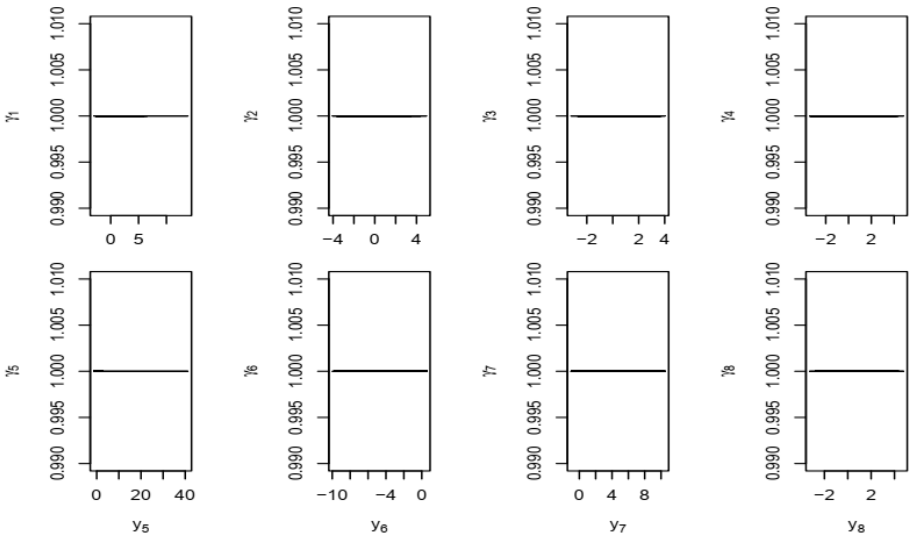

Figure 7: Model (9): Nature of recovery efficiency statistic, γ using Mean weighting scheme

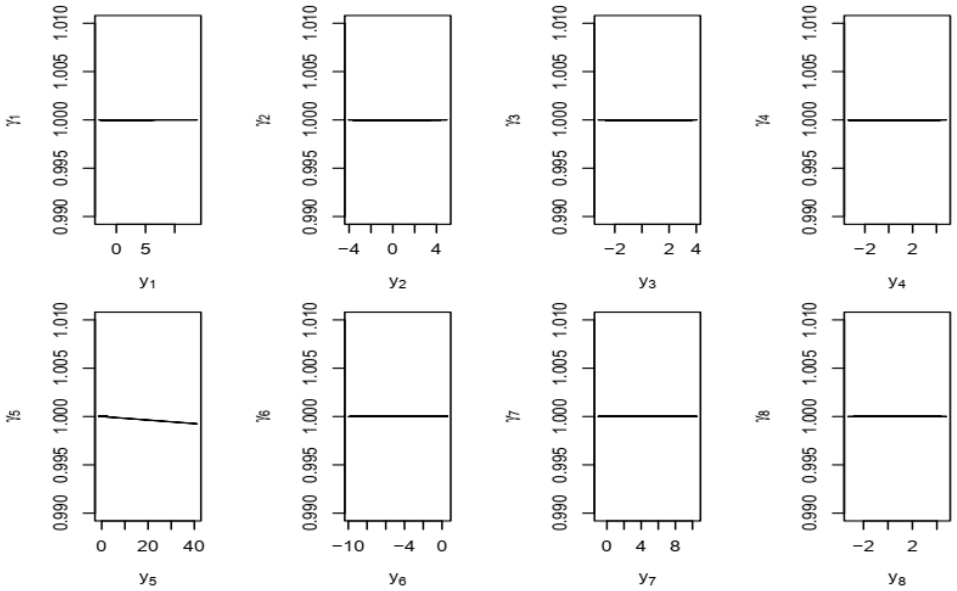

Figure 8: Model (9): Nature of recovery efficiency statistic, γ using Median weighting scheme

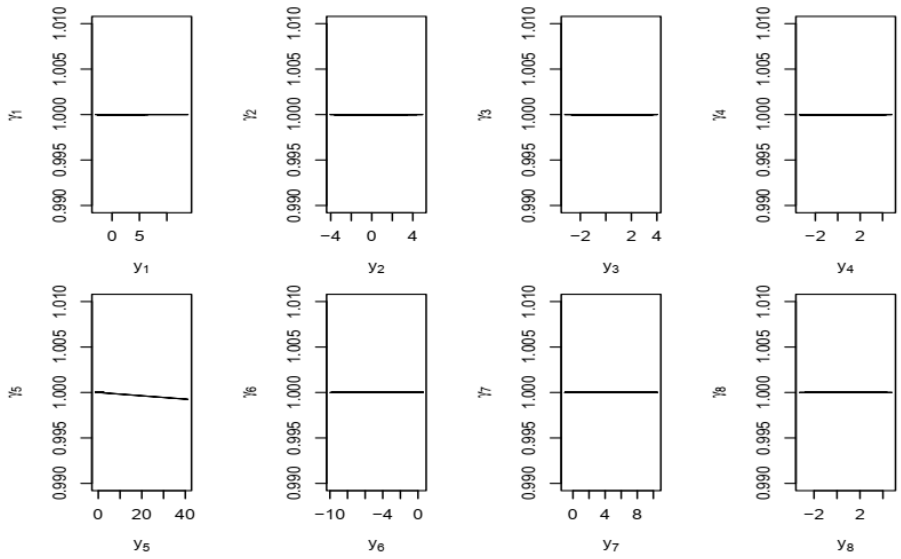

Figure 9: Model (9): Nature of recovery efficiency statistic, γ using OGK weighting scheme

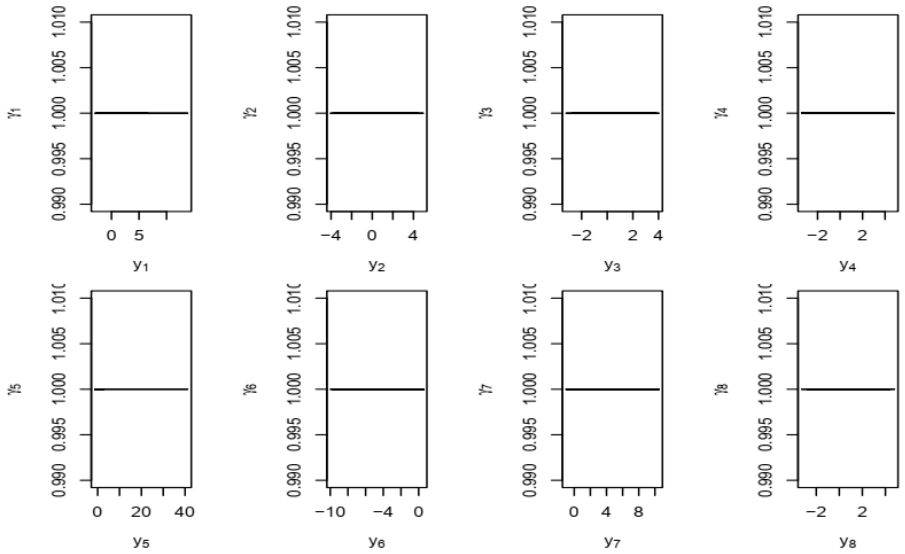

Figure 10: Model (13): Nature of recovery efficiency statistic, γ using Mean weighting scheme

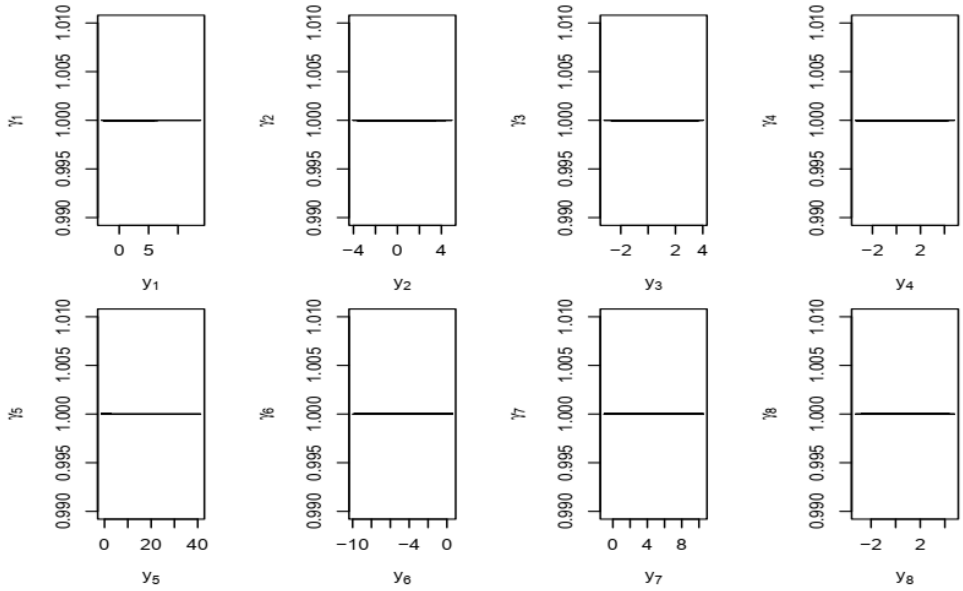

Figure 11: Model (13): Nature of recovery efficiency statistic, γ using Median weighting scheme

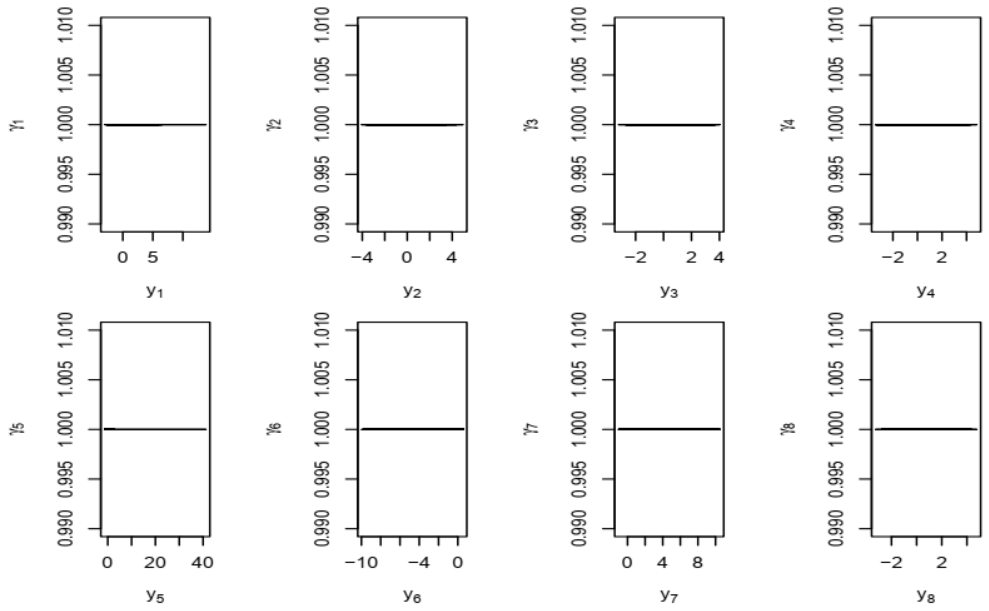

Figure 12: Model (13): Nature of recovery efficiency statistic, γ using OGK weighting scheme

Table 2: Computational Time for Model (9) and Model (13)

	Fusion weight					
	Model (9)			Model (13)		
Error Statistics	t (sec)	t (sec)	t (sec)	t (sec)	t (sec)	t (sec)
Mean	0.00	0.00	0.00	0.00	0.00	0.00
Median	0.00	0.00	0.00	0.00	0.00	0.00
OGK	0.00	0.00	0.00	0.00	0.00	0.00

The proposed method demonstrates a strong ability to recover fused multiple physiological vital signs from a single composite measure. This is evidenced by the low error statistics (RMSRE, RMARE, RSMARE) for the mean, median, and OGK weighting methods shown in Table 1. The fact that these three weighting methods produce identical error statistics suggests a robust recovery process that is not overly sensitive to the choice of weighting method. Again, Figures 1-6 visually confirm the high-quality recovery of individual vital signs from the fused composite measure. The recovered data (red lines) closely follow the original data (black lines) for all eight vital signs, regardless of the weighting method used (mean, median, or OGK). This demonstrates the method's ability to preserve essential information from each vital sign during the recovery process.

Also, the similar performance of mean, median, and OGK weighting methods suggests that the recovery approach is flexible and can adapt to different data characteristics. This could be particularly valuable when dealing with physiological data that may have outliers or non-normal distributions. Lastly, the proposed approximate recovery model is computationally efficient while still providing accurate results. This could be advantageous in real-time monitoring situations where quick data processing is crucial. The recovery method preserves both long-term trends and short-term fluctuations in the vital signs. This is critical for maintaining the clinical relevance of the recovered data. Also, dimensionality is crucial when dealing with multivariate data. By successfully fusing multiple vital signs into a single composite measure, this method offers a way to reduce the dimensionality of complex physiological datasets without significant loss of information.

4. CONCLUSION

The empirical data recovery method for composite similarity-based fusion presented in this study demonstrates a promising approach for handling multivariate physiological vital sign data. The method effectively recovers fused multiple vital signs that is in a single composite measure. It really allows for accurate recovery of individual vital signs from the fused measure, preserving important clinical information. In addition, it shows robustness across different weighting methods (mean, median, OGK), suggesting broad applicability to various types of physiological data. The method's ability to switch between multivariate and univariate representations offers flexibility in data analysis and modelling approaches. These results suggest that this recovery method could have valuable applications in clinical settings, particularly in areas such as patient monitoring, early warning systems, and physiological modelling. The ability to work with a simplified, fused representation of multiple vital signs while retaining the option to recover individual signals could streamline data processing and analysis in healthcare applications.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interest.

ACKNOWLEDGEMENTS

The first author wants to thank the African Union through the Pan African University Institute for Basic Sciences, Technology, and Innovation (Pan African University Scholarship) for their support in the form of Scholarship for my studies.

REFERENCES

[1] S. Wang, M.E. Celebi, Y.D. Zhang, et al. Advances in Data Preprocessing for Biomedical Data Fusion: An Overview of the Methods, Challenges, and Prospects, Inf. Fusion 76 (2021), 376–421. https://doi.org/10.1016/j.inffus.2021.07.001.

[2] M.D. Santos, Vital-Sign Data-Fusion Methods to Identify Patient Deterioration in the Emergency Department, Thesis, University of Oxford, 2018.

[3] M.A.F. Pimentel, D.A. Clifton, L. Clifton, Vital-Sign Data Fusion Models for Post-Operative Patients, in: Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing, SciTePress, Vilamoura, Algarve, Portugal, 2012: pp. 410–413. https://doi.org/10.5220/0003789104100413.

[4] F. Eyiah-Bediako, D.K. Mensah, S. Assabil, et al. Data Fusion for Physiological Vital Signs with Automatic Extreme Value Control, J. Biomed. Eng. 40 (2023), 218-235.

[5] S. Hassani, U. Dackermann, M. Mousavi, J. Li, A Systematic Review of Data Fusion Techniques for Optimized Structural Health Monitoring, Inf. Fusion 103 (2024), 102136. https://doi.org/10.1016/j.inffus.2023.102136.

[6] S. Khalid, Data Fusion Models for Detection of Vital-Sign Deterioration in Acutely Ill Patients, Thesis, University of Oxford, 2013.

[7] Z. Liu, H.A. Song, V. Zadorozhny, et al. H-Fuse: Efficient Fusion of Aggregated Historical Data, in: Proceedings of the 2017 SIAM International Conference on Data Mining, SIAM, Philadelphia, PA, 2017. https://doi.org/10.1137/1.9781611974973.

[8] Y. Zhao, X. Zhang, J. Wang, et al. A New Data Fusion Driven-Sparse Representation Learning Method for Bearing Intelligent Diagnosis in Small and Unbalanced Samples, Eng. Appl. Artif. Intell. 117 (2023), 105513. https://doi.org/10.1016/j.engappai.2022.105513.

[9] R.C. King, E. Villeneuve, R.J. White, et al. Application of Data Fusion Techniques and Technologies for Wearable Health Monitoring, Med. Eng. Phys. 42 (2017), 1–12. https://doi.org/10.1016/j.medengphy.2016.12.011.

[10] H. Lee, K. Park, B. Lee, et al. Issues in Data Fusion for Healthcare Monitoring, in: Proceedings of the 1st International Conference on PErvasive Technologies Related to Assistive Environments, ACM, Athens Greece, 2008: pp. 1–8. https://doi.org/10.1145/1389586.1389590.

[11] R.T. Wu, M.R. Jahanshahi, Data Fusion Approaches for Structural Health Monitoring and System Identification: Past, Present, and Future, Struct. Health Monit. 19 (2020), 552–586. https://doi.org/10.1177/1475921718798769.

[12] A. Jazayeri, O.S. Liang, C.C. Yang, Imputation of Missing Data in Electronic Health Records Based on Patients' Similarities, J. Healthc. Inform. Res. 4 (2020), 295–307. https://doi.org/10.1007/s41666-020-00073-5.

[13] H. Khan, M.T. Rasheed, H. Liu, S. Zhang, High-Order Polynomial Interpolation with CNN: A Robust Approach for Missing Data Imputation, Comput. Electr. Eng. 119 (2024), 109524. https://doi.org/10.1016/j.compeleceng.2024.109524.

[14] D. Labate, F.L. Foresta, G. Occhiuto, et al. Empirical Mode Decomposition vs. Wavelet Decomposition for the Extraction of Respiratory Signal from Single-Channel ECG: A Comparison, IEEE Sens. J. 13 (2013), 2666–2674. https://doi.org/10.1109/JSEN.2013.2257742.

[15] P. Nguyen, J.M. Kim, Adaptive ECG Denoising Using Genetic Algorithm-Based Thresholding and Ensemble Empirical Mode Decomposition, Inf. Sci. 373 (2016), 499–511. https://doi.org/10.1016/j.ins.2016.09.033.

[16] M.A. Ofori, D.K. Mensah, G.O. Orwa, P. Hewson, Traumatic Physiological Vital Sign Fusion: Insight from Composite Spatial Similarity Measure Modelling, TechRxiv (2024). https://doi.org/10.36227/techrxiv.173202834.44153556/v1.

[17] C.E. Rasmussen, Gaussian Processes in Machine Learning, in: O. Bousquet, U. Von Luxburg, G. Rätsch (Eds.), Advanced Lectures on Machine Learning, Springer Berlin Heidelberg, Berlin, Heidelberg, 2004: pp. 63–71. https://doi.org/10.1007/978-3-540-28650-9_4.

[18] R.A. Maronna, R.D. Martin, V.J. Yohai, Robust Statistics: Theory and Methods, Wiley, 2006. https://doi.org/10.1002/0470010940.

[19] D.K. Mensah, M.A. Ofori, N. Howard, Traumatic Systolic Blood Pressure Modeling: A Spectral Gaussian Process Regression Approach with Robust Sample Covariates, Math. Stat. 10 (2022), 562–574. https://doi.org/10.13189/ms.2022.100312.