Impact of similarity measure and fusion statistics on vital signs data fusion

Micheal Arthur Ofori[1*], David Kwamena Mensah[2], George Otieno Orwa[3], and Paul Hewson[4]

**E-mail: arthur.michael@students.jkuat.ac.ke*
ORCID: 0000-0002-4983-540X

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

KEYWORDS

vital signs monitoring, data fusion, similarity measures, fusion statistics, nonlinear analysis, physiological measurements

ABSTRACT

This study investigates the impact of different similarity measures and fusion statistics on vital signs data fusion, comparing the performance of linear and nonlinear methodologies. We compared kappa (κ) and tau (τ) as nonlinear measures against Pearson correlation (ρ) as linear measures, with different statistical approaches (mean, median, and Orthogonalized Gnanadesikan-Kettenring (OGK)). Performance was assessed using grand mean absence statistics ($\bar{R}$) across four competing models with different combinations of original observations (y) and transformed data (d(y)) for parameter estimation (θ) and fusion statistics (g(y)). The results demonstrate that nonlinear measures consistently outperform linear measures, with κ providing higher $\bar{R}$ estimates than τ across all statistical approaches. For nonlinear measures, using original observations for parameter estimation and transformed data for fusion statistics consistently yielded superior results. Linear measures showed more variability in performance depending on the specific measure and statistical approach used. The robust statistics (median and OGK) generally provided better performance than mean statistics, likely due to their resistance to outliers. This methodology offers significant practical utility for public health applications, including detection of malfunctioning organs, design of clinical tests, comprehensive health monitoring, and automated problematic vital organ identification. The approach demonstrates particular promise for conditions like diabetes, where current diagnostic methods rely on isolated measurements rather than integrated vital sign assessment.

Introduction

Physiological vital signs play a crucial role in emergency health and have become an integral part of healthcare research. Medical research exclusively focuses on vital signs fusion because it helps in improving patient monitoring and identifying clinical deterioration early. The health condition of a patient becomes evident through measurements of heart rate along with respiratory rate, blood

pressure, body temperature, and other vital signs. Various data fusion techniques enable the integration of vital signs parameters which elevate health assessment accuracy and reliability for faster interventions and superior patient results (Charlton et al., 2016; Sun et al., 2018). Recent research indicates that optimization of data fusion requires proper implementation of similarity measures and fusion statistics. The extraction and optimization of vital sign information from various data sources become more effective when healthcare professionals apply Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks (Yang et al., 2021; Liu et al., 2019).

These techniques not only increase the accuracy of vital sign monitoring but also make it easier to incorporate extra contextual data, like patient activity and ambient conditions, which can enhance the monitoring systems' predictive power (Yang et al., 2021). The quality standards of vital sign information determine how well clinical decision support systems operate. The quality of vital sign information including accuracy and completeness and timeliness influences both the generation of valid triage scores and clinical choice solutions (Skyttberg et al., 2016). Robust data fusion methods allow healthcare providers to enhance the quality of their patient care data while handling missing data situations (Rossum et al., 2023).

The methods for similarity measure along with fusion criteria demonstrate substantial influence on vital signs data fusion which impacts both clinical operational aspects and the technical phase of processing. To improve patient outcomes and advance the area of medical informatics, it will be crucial to comprehend and optimize these characteristics as healthcare continues to shift towards more data-driven approaches (Sadasivuni et al., 2021).

Wang et al. (2021) developed the basic concept for composite similarity measurement by combining several features to enhance fusion precision. Multi-feature approaches in their research achieved outstanding outcomes which surpassed traditional single-feature methods. Ofori et al (2024) developed a fusion method using composite spatial similarity measure modelling. Their framework was built on mixture random variate using information provided by the interrelationships among variables. That allowed the latent unidimensional data to be generated as a weighted linear combination of the multivariate data, providing an easy way to model the weights in terms of desirable data features of interest.

The current paper describes a new empirical method to evaluate how similarity measures together with fusion statistics influence vital signs data fusion. This method was developed solely for composite similarity assessments in physiological vital signs fusion operations. The algorithm utilizes built-in relationships between vital signs together with characteristics and measurement uncertainties.

Materials and Methods

Data

The data example here is based on the traumatic vital sign data employed by Mensah et. al. (2024). The data was sourced from the Komfo Anokye Teaching Hospital (KTH). We used a de-identified subset characterized by variables RR, HR, SBP, DBP, TEMP, SPO2, RBS, and MAP, of dimension 4064×8.

Model

Physiological vital sign data fusion focuses on having single random variable for p corrected set of variables generated from the same source so that a single model can be developed instead of multiple dependent models. Given p random variables with interrelationships say, $y_1, y_2, \ldots, y_p$, a typical data fusion model according to Ofori et al., (2024), assumes the form

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

for fusion weights θ_{ij} and fusion statistics $g\left(y_{ij}\right)$, where $\sum_{j=1}^{q} \theta_{ij} = 1$ for distributional validity of the fused variable $\tilde{y}$. The choice of θ and $g\left(y_{ij}\right)$ are crucial in data fusion as each comes with its impact on the resultant one-dimensional data pattern. It is thus better to explore varied types specifications spanning linear and non-linear similarity measure-based statistics for θ and fusion statistics with automatic self-controlled and non-controlled data issues to assess their impact on data fusion. We consider the following similarity measure functions for estimating θ

$$k\left(a_1,a_2,a_3,b_1,b_2,b_3,\delta\right) = k_1\left(a_1,b_1,\delta\right) + k_2\left(a_2,b_2,\delta\right) + k_3\left(a_3,b_3,\delta\right), \qquad (2)$$

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

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

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

$$a_1,a_2,a_3,b_1,b_2,b_3 > 0$$

$$\tau\left(a_1,a_2,a_3,b_1,b_2,b_3,\delta\right) = \tau_1\left(a_1,b_1,\delta\right) + \tau_2\left(a_2,b_2,\delta\right) + \tau_3\left(a_3,b_3,\delta\right), \qquad (3)$$

$$\tau_1\left(a_1,a_2,b_1,\delta\right) = \sigma_1^2 \exp\left(-\left(\frac{\delta}{b_1}\right)^2\right) + a_2^2 v$$

$$\tau_2\left(a_2,b_1,b_2,\delta\right) = \sigma_1^2 \exp\left(-\left(\frac{\delta^2}{2b_1^2}\right)-\left(\frac{1-\cos(2\pi\delta)}{b_2^2}\right)\right)$$

$$k_3\left(a_3,b_2,b_3,\delta\right) = \sigma_3^2\left(1+\frac{\delta^2}{2b_2b_3}\right)^{-b_3},$$

where $\delta = \delta\left(y_i,y_j\right) = \left|y_i - y_j\right|$, v is an $n \times n$ matrix with entries computed using

$$v_{ij} = \begin{cases} 1 & \text{if } y_i = y_j \\ 0 & \text{otherwise} \end{cases}$$

and

$$\rho\left(y_j\right) = \frac{\left(y_j - \bar{y}_j\right)}{\sqrt{\sum_{i=1}^{n}\left(y_j - \bar{y}_j\right)^2}} \qquad (4)$$

The actual empirical computation of θ_{ij} based on (2) and (3) following Ofori et al. (2024) considers the mean, median, and Orthogonalized Gnanadesikan Ketterning (OGK) (Maronna et al., 2019, Mensah et al., 2022) defined as

$$\theta_{11} = \frac{1}{n}\sum_{g=1}^{n} k\left(a_1,a_2,a_3,b_1,b_2,b_3,\delta_{1l}\right),\ \delta_{1g} = \delta\left(y_{11},y_{1g-1}\right) \qquad (5)$$

By (5), a corresponding OGK θ_{11} is computed as

$$\theta_{11} = \frac{\sum_{j=1}^{n} \theta_{1j} v\left(z_{\theta_{1j}}\right)}{\sum_{j=1}^{n} v\left(z_{\theta_{1j}}\right)}, \; z_{\theta_{1j}} = \frac{\theta_{1j} - \tilde{\mu}_0}{\tilde{\sigma}_0} \tag{6}$$

$$v\left(z_{\theta_{1j}}\right) = \left[1 - T^2\left(z_{\theta_{1j}}, d\right)\right]^2 I_{\left(\left|z_{\theta_{1j}}\right| \le d\right)}, \; T^2\left(z_{\theta_{1j}}, d\right) = \frac{z_{\theta_{1j}}}{d} \tag{7}$$

where $b = 4.5$, $\tilde{\mu}_0$ and $\tilde{\sigma}_0$ give the median and absolute deviation (MAD) of θ_{1j} respectively. A corresponding estimate of θ suing (4) is computed as

$$\theta_{11} = \frac{1}{m} \sum_{l=1}^{m} S_r\left(g_{(y_{11})}, g_{(y_{1l})}\right) \tag{8}$$

Choice of statistics for θ and $g(y)$

We consider the choice of potential data and statistics for computing θ and $g(y)$ respectively, as their choice can influence the resultant data pattern both positively and negatively. Importantly, both have the potential for handling hidden data issues that may affect model building at the variable level. As such, it must be controlled automatically to avoid unnecessary drastic effects in the combined form. In particular, we consider the observation, y, and its feature counterpart, $d(y)$ so that $g(y)$ can be set as $g(y) = y$ and $g(y) = d(y)$. With this θ can be computed using y and $d(y)$. Thus, it is easy to see the possible fusion models that can be generated from the possible combinations of the resulting θs based on the three similarity measures (2), (3) and (4) and $g(y)$s. In what follows, we provide detail on the derivation of $d(y)$ statistics. Moments of random variables provide viable sources of features for handling hidden data issues such as extreme, repeated values based on their natural function of correct weighting information. Let y generate probability density function $f(y)$. Then, its t moment about the origin is computed as

$$E\left[Y^k\right] = \int y^t f(y) dy \tag{9}$$

Obviously, the data for computing $E\left[Y^k\right]$ has the form $y^t f(y)$. Thus, this can be viewed as a feature of y under the probability density space since there is a direct link with the y in terms of recovery (Mensah et al., 2022). Based on above information, we set $d(y)$ as

$$d(y) = \sqrt{y^t f(y)}. \tag{10}$$

The underlying probability density $f(y)$ computed using the kernel density estimation implemented in R (Scott, 2015, Silverman, 1986). Table1 provides a comprehensive summary of possible models that results from the above data for computation of θ and statistics for use as fusion statistic, $g(y)$.

Table 1: Candidate Competing Models

Data for θ and statistics for $g(y)$		
Similarity Measure	θ	$g(y)$
κ	y, $d(y)$	y, $d(y)$
τ	y, $d(y)$	y, $d(y)$
S_r	y, $d(y)$	y, $d(y)$

The computation of estimates for parameters $a_1, a_2, a_3, b_1, b_2, b_3$ follow Ofori et al., (2024).

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

where we define $\gamma(1), \gamma(2), \ldots, \gamma(n)$ are the order statistics of γ, and $\overline{d(y)}$ and $\tilde{d}(y)$ represent the mean and median of $d(y)$ respectively.

We outline the channels for automatic control of data issues, key aspects of the similarity measures bring the novelty of specifications here. Typical automatic control of data issues functionality the similarity measures are captured in δ for (2) and (3) and μ for (4) while in the case of ρ, it is induced in the differences between competing observations $\left(y_i - y_j\right)$. Thus, all levels of point-auto-relationships are captured with high weight for near points and low weights for far ones. Finally, the usage of $d(y)$ as input data for δ and ρ ensures the multiple automatic control of data issues iterated in this paper. It becomes straightforward to appreciate the levels of automatic controls that the above data combinations offer in fusion. For example, (y, y) with any method for computing θ say mean, median and OGK will generate the least number of automatic controls based on any similarity measure.

Vital sign absence relevance statistics

We analyze the impact of similarity measure and choice of fusion statistics via vital sign relevance in $\tilde{y}$ based on measures. Consider the recovery model for $g\left(y_j\right)$ obtained by solving (1)

$$g\left(y_j\right) = \theta_j^{-1}\tilde{y} - \sum_{l=2}^{q} \lambda_{jl} g\left(y_l\right), \quad \lambda_{jl} = \frac{\theta_l}{\theta_j} \qquad (11)$$

If we let $M\left(y_j\right) = \sum_{l=2}^{q} \lambda_{jl} g\left(y_l\right)$, then by (11), it can be seen that the statistic $M\left(y_{il}\right)$ Provides insight into how much other vital sign variables contribute towards the recovery of variable y_j. Thus, source for an assessment to quantify how much the absence of y_j to the generation of $\tilde{y}$ if all variables operate together. Define $R\left(y_j\right)$ as the mean weighted contribution of all other variables due to the absence of y_j,

$$R\left(y_j\right) = \frac{1}{N}\sum_{l=2}^{q} \lambda_{jl} g\left(y_l\right), \qquad (12)$$

where N denotes the total number of observations in all variables without y_j. A grand statistic based on (12) can be computed and used to conduct impact analysis of the choice of data for computing θ as well as fusion statistic.

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 Fusion Error (RSMAFE) defined for true data y and its recovered counterpart, $\hat{y}$ as follows.

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

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

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

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

$$\gamma_i = \left(\frac{y_i}{\hat{y}_i}\right) \qquad (13)$$

and

$$\bar{\gamma} = \left(\sum_{i=1}^{n}\gamma_i\right)\times 100\% \qquad (14)$$

Results and Discussion

The study evaluated various combinations of fusion statistics and similarity measures to understand their effects on vital signs data fusion by changing weight parameters (θ). The evaluation produced essential results that emerged from different methodological approaches. The empirical values of spatial covariance parameters (Table 2) show significant variation across the eight vital sign variables (y_1 through y_8). Parameter a_1 ranges from 0.2599 to 0.4897, with y_6 showing the highest value (0.4897) and y_5 the lowest (0.2599), indicating varying levels of short-range spatial dependency.

Table 2: Empirical values of spatial covariance parameters

	Empirical Estimates							
Parameter	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
a_1	0.4426	0.3291	0.3217	0.3139	0.2599	0.4897	0.2923	0.3192
a_2	0.4885	0.3538	0.3417	0.3526	0.2557	0.4690	0.2851	0.3418

a_3	0.8125	0.5315	0.5619	0.5509	3.2094	1.6008	1.2660	0.5493
b_1	1.8416	1.6428	1.7644	1.7607	12.3985	3.3410	4.3378	1.7292
b_2	1.8230	1.6428	1.7644	1.7607	1.8009	3.3410	4.3378	1.7292
b_3	1.8230	1.6428	1.7644	1.7607	1.8009	2.9412	4.3378	1.7292

Source: Authors

The empirical values of spatial covariance parameters show significant variation across the eight vital sign variables (y_1 through y_8). Parameter a_1 ranges from 0.2599 to 0.4897, with y_6 showing the highest value (0.4897) and y_5 the lowest (0.2599), indicating varying levels of short-range spatial dependency. The values for parameter a_2 follow comparable patterns while showing somewhat higher numbers from 0.2557 to 0.4885. Parameter a_3 exhibits significant variability through its measured values that extend from 0.5315 to 3.2094 where a high value occurs specifically in y_5 (3.2094). Data shows that the b_1, b_2, b_3 parameters demonstrate consistent tendencies within each variable while displaying significant differences between variables and b_1 for y_5 reaches exceptionally high levels at 12.3985. The b-values of y_6 and y_7 stand out from all other variables' b-values. All variables keep their b-values stable throughout the three-measurement points b_1, b_2, b_3.

In terms of performance of nonlinear similarity measures, throughout various weight calculations κ() and τ() showed identical patterns of performance in vital sign fusion analyses. The performance of nonlinear measures is generally consistent in the grand mean of all absence statistics ($\bar{R}(y)$ across the two measures: kappa (κ()) and tau (τ()) using mean, median, and OGK statistics for all the competing models. Higher $\bar{R}(y)$ values indicate a greater impact on vital sign fusion in terms of efficient calibration of vital signs relevance to a latent underlying one-dimensional pattern that could be viewed as degree of health, if it is of interest to measure health based the joint operation of vital signs. The fusion performance for the mean $(\bar{\theta})$ with κ() measure (Table 3) shows that using the original observation (y) as data for computing θ and d(y) as fusion statistic (g(y)) gives an overall higher $\bar{R}(y)$ (4.0229) among the four possible competing models. This result is consistent with using τ() measure. It also indicates an overall $\bar{R}(y)$ of 3.5837 for using y as data for computing θ and d(y) for fusion statistics. Also, using d(y) for both θ and g(y) has an $\bar{R}(y)$ of 2.3312. This consistency is also seen in all the possible competing models across the median $(\tilde{\theta})$ (Table 4) and OGK (θ_{OGK}) (Table 5). This is followed by the use of d(y) as data to compute both θ and g(y) (2.2792). It can be seen that data specific combinations (y, d(y)), (d(y), y) and (d(y), d(y)) always yield better grand mean absence value. This performance is attributed to the ability to handle within vital sign specific data issues such as extreme values, repeated values, point-wise interrelationships.

According to Table 3 (mean-based weight) for κ() similarity measure, using raw data (y) for both θ and g(y) yields the lowest MSFE (0.0358). Also, feature-based approaches d(y) shows higher error rates when used for θ as MSFE increases to 0.1093. Again, d(y) produces an MSFE of 0.0644 when used for g(y).

Table 3: κ(), τ () : Fusion performance for $\bar{\theta}$ (Mean)

Measure	data for θ	g(y)	M SF E	SM AF F E	M AF E	$\bar{R}(y)$
	Data for θ and statistics for g(y)					
κ()	y	y	0.0358	0.1382	2.3353	-0.75026
	y	d(y)	0.0644	0.2326	0.8750	4.02293
	d(y)	y	0.1093	0.1620	1.3975	-0.15307
	d(y)	d(y)	0.0667	0.2390	0.8750	2.27924
τ ()	y	y	0.0350	0.1367	1.3627	-0.52309
	y	d(y)	0.0644	0.2329	0.8750	3.58370
	d(y)	y	0.1095	0.1616	1.5796	-0.15842
	d(y)	d(y)	0.0906	0.2786	1.0180	2.33124

Source: Authors

Table 4: κ(), τ () : Fusion performance for $\tilde{\theta}$ (Median)

Measure	data for θ	g(y)	M SF E	M AF E	SM AF E	$\bar{R}(y)$
	Data for θ and statistics for g(y)					
κ() κ()	y	y	0.0350	1.5323	8.5066	-427743707
	y	d(y)	0.0646	0.2330	0.8750	7420277938
	d(y)	y	0.1082	0.1624	1.8011	-0.15381
	d(y)	d(y)	0.0667	0.2391	0.875	2.25383
τ () τ ()	y	y	0.0341	0.1360	1.4820	-0.63461
	y	d(y)	0.0646	0.2335	0.8750	3.93213
	d(y)	y	0.1087	0.1619	1.4263	-0.15955
	d(y)	d(y)	0.0910	0.2795	1.0198	2.31745

Source: Authors

According to Table 4 (median-based weights), both similarity measures show improved performance compared to mean-based weights. The κ() measure with raw data achieves MSFE of 0.0350. The τ() measure shows further improvement with MSFE of 0.0341. Also, the feature-based fusion maintains higher error rates but shows more stability. Meanwhile, notable difference exists in $\bar{R}(y)$ values, with extreme values for some combinations. The OGK-based weights (Table 5) show performance metrics that closely mirror median-based weights. There are slight improvements in stability across different combinations with more consistent $\bar{R}(y)$ values

compared to other weight types. It really maintains the pattern of better performance with the raw data fusion.

Table 5: κ(), τ () : Fusion performance information for θ_{OGK}

Measure	data for θ	g(y)	M SF E	M AF E	SM AF E	$R^{-}(y)$
	Data for θ and statistics for g(y)					
κ()	y	y	0.0350	0.1376	1.5323	-407738607
κ()	y	d(y)	0.0646	0.2330	0.8750	7410269113
	d(y)	y	0.1082	0.1625	1.8011	-0.15337
	d(y)	d(y)	0.0916	0.2802	1.0212	2.25355
τ ()	y	y	0.0341	0.1360	1.6027	-0.63334
τ ()	y	d(y)	0.0645	0.2334	0.8750	3.92488
	d(y)	y	0.1086	0.1619	1.4882	-0.15902
	d(y)	d(y)	0.091	0.2795	1.0199	2.31749

Source: Authors

The performance assessment of linear similarity measure on the other hand shows differences in $\bar{R}(y)$ across the two measures considered: Pearson (ρ()) and Euclidean (S_r) using the mean, median and, the OGK for all the possible competing models. Fusion performance for $\bar{\theta}$ (Table 5) with ρ() measure indicates that using the original observations in computing both θ and g(y) provides the best $\bar{R}(y)$ of 3.1907. This is followed by the use of d(y) for both θ and g(y) (0.1307). The story is different from fusion performance with the median (Table 7) and the OGK (Table 8). The median shows that the usage of d(y) for θ and y for computing g(y) provides a better estimate of $\bar{R}(y)$ (0.2559). Meanwhile, the fusion performance with the OGK under the Pearson measure proves that the usage of d(y) for calculating both θ and g(y) is the best $\bar{R}(y)$ estimate compared with all the competing models. The performance of linear similarity measures (Tables 6-8), throughout various weight calculations showed identical patterns of performance in vital sign fusion analyses. The mean-based weight showed significantly different performance patterns from nonlinear measures. Also, the feature-based fusion shows remarkably low error rates with MSFE of 0.0056 for (y, d(y)) combination,

Table 6: Linear similarity measure: Fusion performance for θ with $\bar{y}$ (Mean)

Measure	data for θ	g(y)	M SF E	M AF E	SM AF E	$R^{-}(y)$
	Data for θ and statistics for g(y)					
ρ()	y	y	0.8564	0.1612	0.1612	3.19073

y	d(y)	0.0056	0.0269	0.0269	-0.30289
d(y)	y	0.1617	0.0572	0.0572	0.03966
d(y)	d(y)	0.0023	0.0284	0.0284	0.13072

Source: Authors

MSFE of 0.0023 for (d(y), d(y)) combination. The raw data fusion shows higher error rates (MSFE 0.8564) with $\bar{R}(y)$ values that are more stable and closer to zero. Likewise, the median-based weight (Table 7) exhibited a similar pattern to that of the mean-based weight but with slight improvements. Its feature-based fusion maintains excellent performance with MSFE of 0.0056 for (y, d(y)), MSFE of 0.0020 for (d(y), d(y)) and a more consistent $\bar{R}(y)$ values across all combinations. The OGK-based weight (Table 8) outperforms both the mean and median weight for linear measures. It shows the most stable performance across different combinations. The feature-based fusion continues to show optimal performance with $\bar{R}(y)$ values indicating a better handling of outliers and extreme observations.

Table 7: Linear similarity measure: Fusion performance for θ with $\tilde{y}$ (Median)

Data for θ and statistics for g(y)						
Measure	data for θ	g(y)	M SF E	M AF E	SM AF E	$\bar{R}(y)$
ρ()	y	y	0.8537	0.1589	0.1589	-0.25440
	y	d(y)	0.0056	0.0261	0.0261	-0.01928
	d(y)	y	0.1608	0.0502	0.0502	0.25593
	d(y)	d(y)	0.0020	0.0246	0.0246	0.01219

Source: Authors

Table 8: Linear similarity measures: Fusion performance for y_{OGK}

Data for θ and statistics for g(y)						
Measure	data for θ	g(y)	M SF E	M AF E	SM AF E	$\bar{R}(y)$
ρ()	y	y	0.8541	0.1587	0.1587	-1.23668
	y	d(y)	0.0056	0.0260	0.026	-0.51719
	d(y)	y	0.1611	0.0527	0.0527	-0.22091
	d(y)	d(y)	0.0021	0.0252	0.0252	0.33651

Source: Authors

Both similarity measures reported Minimal Mean Square Fusion Error values from Tables 3-5 especially when one uses the raw data (y) for weight calculation and fusion statistics conditions. The MSFE values remained within the range of 0.0341 to 0.0358 showing effective fusion accuracy. The use of derived statistics d(y) resulted in significant increases in error measures when utilized for this specific similarity measures. Raw data therefore appears more appropriate for these measures. The error pattern analysis shows that the Standardized Mean Absolute Fusion Error (SMAFE) revealed specific patterns among the different tested combinations. The SMAFE

measurement levels remained higher within both raw data-based fusion statistics and all weight calculation methods for nonlinear evaluation entities. Standardization influences fusion outcomes to a stronger extent during nonlinear data processing that uses raw data measurements.

Figures 1-3 show the absence relevance statistics using the Pearson, κ(), and τ() measures. The Pearson measure (Figure 1) shows the most consistent patterns across variables with a clear differentiation between different fusion approaches. It also shows stable behavior across all eight variables. The κ() exhibits higher variability in R(y) values with some extreme values indicating potential instability. It also reveals clear differences between raw and feature-based fusion. Lastly, the τ() measures shows average stability between Pearson and κ measures with more consistent patterns than κ() but less stable than Pearson. It reveals a clear impact of different fusion approaches

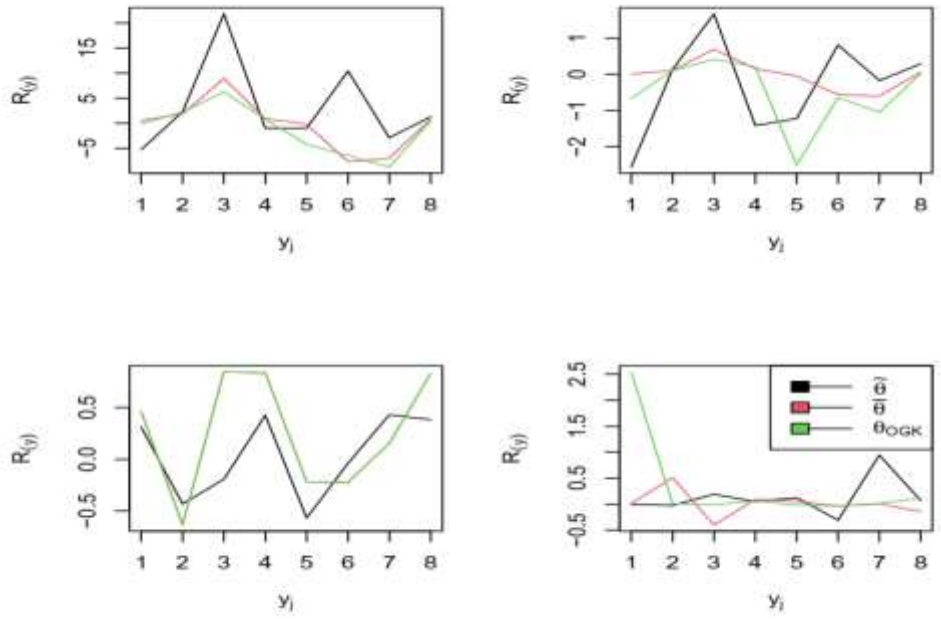

Figure 1: Nature of R(y_j) statistics based on Pearson similarity measure. From left to right are plots for (y, y), (y, d(y)), (d(y), y) and (d(y), d(y)) respectively.

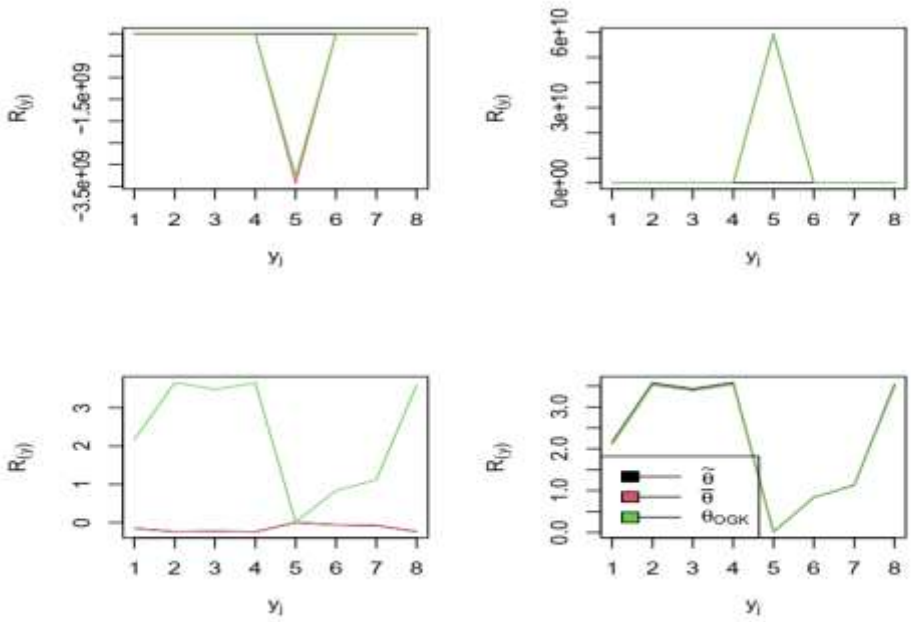

Figure 2: Nature of $R(y_j)$ statistics for κ() similarity measure. From left to right are plots for (y, y), (y, d(y)), (d(y), y) and (d(y), d(y)) respectively.

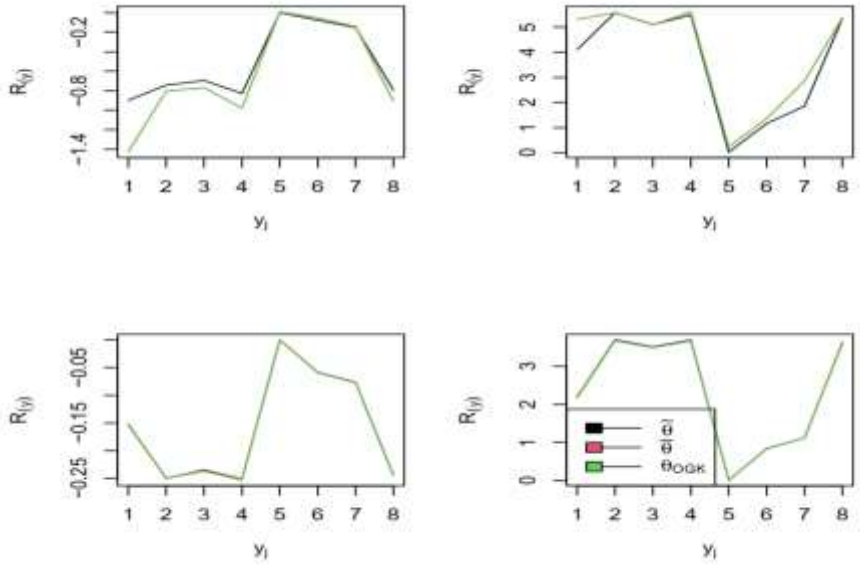

Figure 3: Nature $R(y_j)$ statistics for τ() similarity measure. From left to right are plots for (y, y), (y, d(y)), (d(y), y) and (d(y), d(y)) respectively.

The visual representations of the R(y) statistics in Figures 1-3 demonstrate the influence that the absence of individual vital signs has on fusion processes. Different vital signs exhibit different degrees of sensitivity according to the patterns while some vital signs demonstrate stronger responses than others. The nonlinear measures exhibited better stability for R(y) patterns than the linear measure when working with derived statistics.

Figures 4-6 present visual evidence which demonstrates that κ() and τ() show uniform performance patterns during different weight calculation procedures. The case is somehow different from the

evidence in Figure 7. These plots show minimal discrepancy between the two measures and stable fusion patterns, especially when raw data is used.

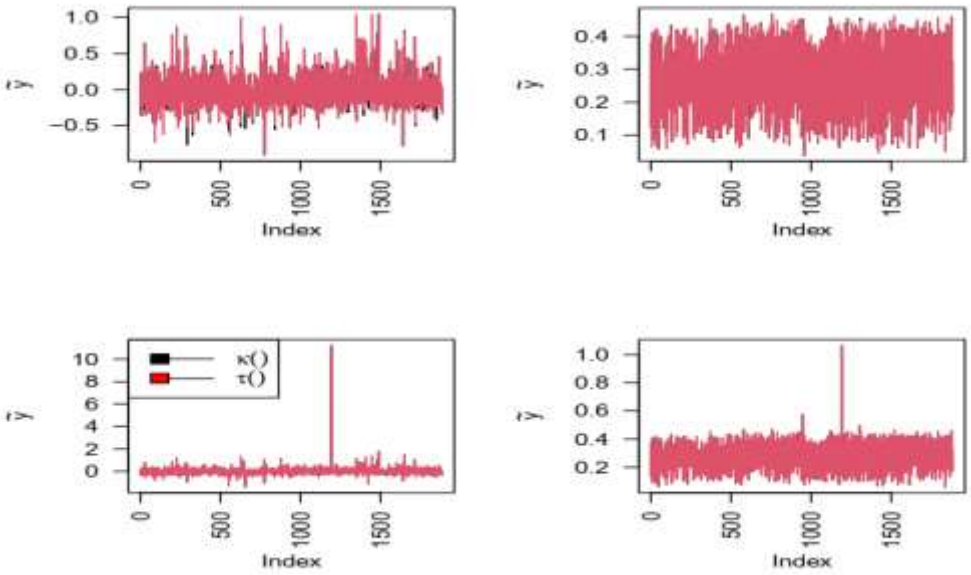

Figure 4: Pattern of $\tilde{y}$ using $\kappa()$ – black and $\tau()$ – Red with $\bar{\theta}$. From left to right are plots for (y, y), (y, d(y)), (d(y), y) and (d(y), d(y)) respectively

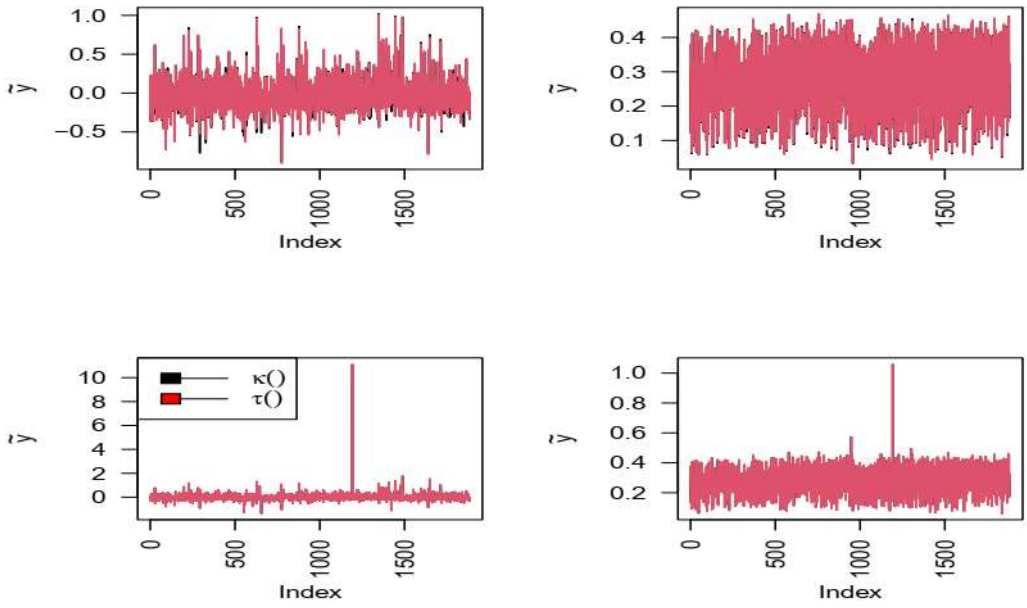

Figure 5: Pattern of $\tilde{y}$ using $\kappa()$ – black and $\tau()$ – Red with $\tilde{\theta}$. From left to right are plots for (y, y), (y, d(y)), (d(y), y) and (d(y), d(y)) respectively

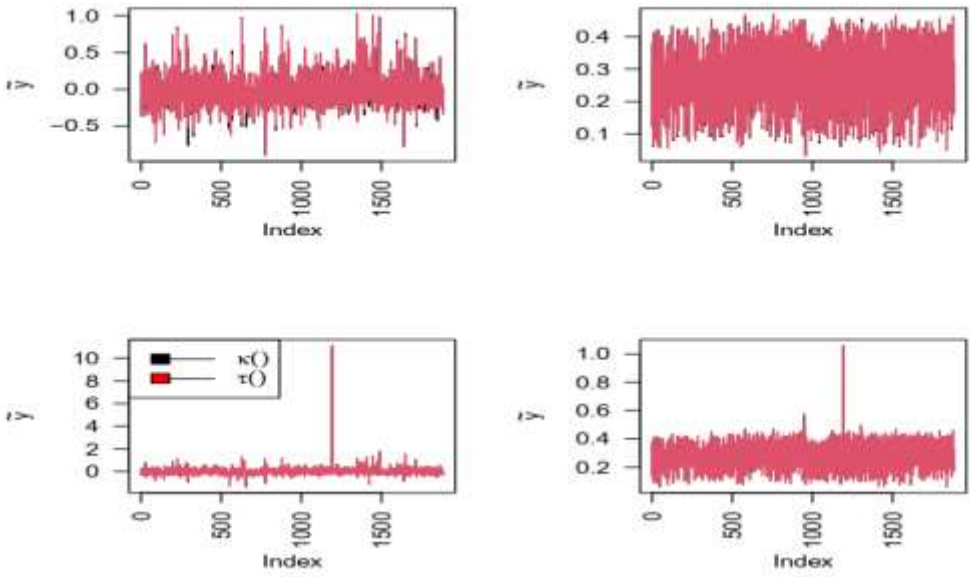

Figure 6: Pattern of $\tilde{y}$ using $\kappa()$ – black and $\tau()$ – Red with θ_{OGK}. From left to right are plots for

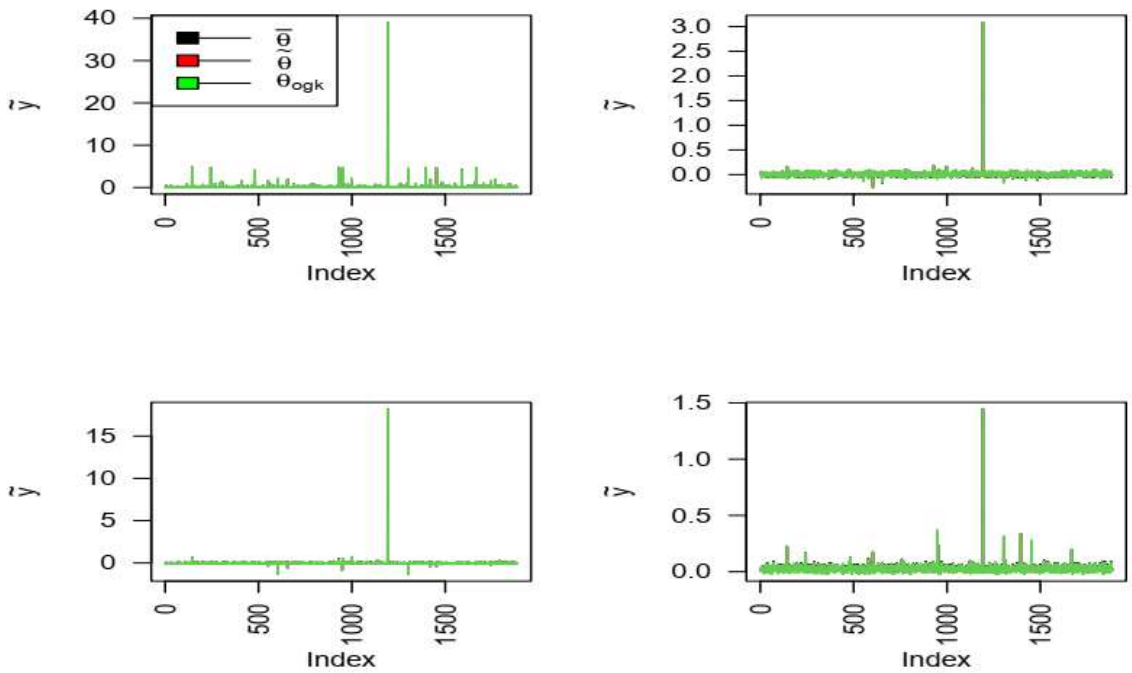

Figure 7: Pattern of $\tilde{y}$ using $\rho()$. From left to right are plots for (y, y), $(y, d(y))$, $(d(y), y)$ and $(d(y), d(y))$ respectively.

(y, y), (y, d(y)), (d(y), y) and (d(y), d(y)) respectively. Comparing the two non-linear measures, it is obvious that the κ() provides higher estimates of $\bar{R}(y)$ than τ() across mean, median, and OGK statistics. The choice between mean, median, and OGK for non-linear measures is also apparent. The estimates $\bar{R}(y)$ from the mean are lower compared to the median and the OGK.

This might be as a result of the mean's inability to deal with extreme observations and outliers. Also, using the original observations for estimating θ and d(y) for the fusion statistic yields higher results in $\bar{R}(y)$. The linear similarity measure on the other hand is very diverse. The choice of fusion performance is dependent on the type of linear measure, the choice of statistic (mean, median, and OGK) as well as the competing model. Fusion performance with mean under the Pearson measure indicates that the use of y for computing both θ and g(y) is the ideal competing model to provide significant $\bar{R}(y)$ estimates. In terms of the OGK, it is better to consider the d(y), d(y) competing models. Likewise, the use of d(y) for θ and y for fusion statistics provides better $\bar{R}(y)$ estimates with the median. Lastly, the choice of Euclidean measure provides two competing models (y, d(y)) and (d(y), d(y)) to provide significant estimates of $\bar{R}(y)$ irrespective of the statistic used (mean, median, or OGK). This implies that the kind of data preparation and the particular needs of the fusion task should have an impact on the choice of similarity measure. The parameter estimation results in Table 2 demonstrate distinctive values for each vital sign which demonstrates that the fusion mechanism adjusts to the respective features of individual vital sign variables. The ability of this system to adjust through different physiological measurements serves as a vital component for achieving accuracy in fusion processes. The current findings suggest that the choice of similarity measure and fusion statistics should be carefully considered based on the specific characteristics of the vital sign data and the intended application. The results indicate that no single approach consistently outperforms the others across all scenarios, highlighting the importance of context-specific selection of fusion methods.

Discussion

Similarity measures and fusion statistics are important in vital signs data fusion to determine the degree of success in the integration and interpretation of physiological fused data. The vital sign specific R(y) statistics values can be viewed as a path for evaluating the effectiveness of a given similarity measure in calibrating the impact a vital sign's absence creates in the ecosystem span by the vital signs. Figures 1-4 show the absence relevance statistics path for the eight vital signs using the κ(), and τ() measures etc. The visible dynamic differences in pattern exhibited by vital sign specific absence statistics across the similarity measures and four fusion data configurations (y, d(y)), (d(y), d(y), d(y)) is a key indicator of flexibility and innovation potential that the new area of data fusion method can offer to consolidated handling of public health issues.

The selection of similarity measures especially the non-linear ones has been proven to improve the fusion process by capturing relatively complex relationships that exists in physiological data. For instance, the application of the τ() measure helped to make a transition in fused visuals more natural, and the median estimation approach helps to retain outlier details which are mentioned by researchers in earlier studies (Dziorny et al., 2022; Gao et al., 2018). Among all the estimation methods, OGK becomes the most suitable approach, stabilizing both direction changes and outliers as key to maintaining effective clinical care and patient outcomes in the case of displaying significant big data outliers (Jesus & Silva, 2022).

Furthermore, the d(y) fusion statistics which form part of the data fusion framework was another improvement in the data fusion process. Besides increasing the representation of vital-sign relationships, this approach also proves consistent in optimizing the management of hidden data problems. Natural interactions happen at the physiological level and traditional statistics are not sensitive enough to capture these differences; that is why d(y) statistics were developed (Zhang et al., 2015). The use of d(y) statistics along with y-based θ shows that while providing a good

balance between signal features and noise reduction, the type of method used impacts the effectiveness of data fusion (Habib et al., 2017).

The complete configuration employing both d(y) statistics and d(y)-based θ estimates proved to be the most complex based on the results obtained in this study. Thus, this method produces the highest complexity of the pattern structures and shows maximum responsiveness to the features of the underlying data. The comparison between τ() and κ() measures further increases, as τ() yields a stable pattern while κ() increases the sensitivity to local details (Yan et al., 2017). However, as the above analysis indicated, linear similarity measures, although fewer in computation, have their own merits in terms of computational simplicity and straightforward interpretation and hence can be applied where computation complexity is not essential (Liu et al., 2016).

The comparisons of the linear and non-linear methods disclose that the latter provide higher accuracy in comparison with the linear ones used for describing intensive physiological interactions. However, linear measures are useful because they are easy to interpret and compute, which can be beneficial in continuous assessments of datasets stocking (Peng et al., 2022). The choice of fusion statistics is also crucial; although y-based fusion has a straightforward relation to the actual measure and does not require the complexity of d(y), d(y)-based fusion deals with the considered hidden data problem and outliers, pointing to simplicity-robustness trade-offs (Hu et al., 2021).

For practical applications, these findings suggest different approaches based on specific needs. Complex vital sign monitoring scenarios would benefit most from non-linear similarity measures combined with d(y) fusion statistics and OGK estimation. Routine monitoring might be better served by linear similarity measures with y-based fusion statistics and median estimation. Research applications might require implementing multiple approaches to compare results and select the most appropriate method based on specific objectives. The collective analysis of these results underscores the importance of careful method selection in vital signs data fusion. The choice of approach can significantly impact the quality and utility of the fused data, making this understanding crucial for both clinical applications and research purposes. As we continue to advance in this field, these insights will prove valuable in developing more refined and effective fusion techniques for vital signs monitoring and analysis.

Conclusion

The study reveals fascinating proof that the choice of similarity measures and fusion statistics directly determines the effectiveness of vital signs data fusion. The results show that when applied to raw vital signs data, nonlinear similarity measures (κ and τ) perform exceptionally well. Also, on the other hand, the linear similarity measure (ρ) produces exceptionally low MSFE values of when used with derived statistics. These results demonstrate that no single strategy consistently performs better than others in every situation, highlighting the significance of choosing a methodology that is appropriate for the given environment.

Generally, whiles it is different to provide a simplistic view of the impact assessment with use of fusion error regardless of nature of similarity measures, it is pretty easy to achieve that with the grand absence statistic $\bar{R}(y)$. This is because the variation in fusion error values across the similarity measures for the four combinations (y, y), (y, d(y)), (d(y), y) and (d(y), d(y)) exhibits inconsistencies. These inconsistencies may suggest that impact assessment should not only be based on the error incurred in generating a composite vital sign from the multivariate vital sign data since there may be other hidden factors that may have contributed to such variation.

Thus, there should be another statistic to help calibrate how the absence of a given vital sign affects the operation of the remaining since they work in tandem (collectively) to determine the health of

an individual at any given time. Obviously, such statistics will deliver great benefits in several ways, for example, directing innovative methods for development of devices for detection of malfunctioning organs, effective design of clinical test for assessment of diseases, wholistic health monitoring via composite deterioration detection, automated problematic vital organ identification and automated self-healing vital organs device design. One practical disease area where this innovation can be of great benefit is the detection and management of diabetes, which is based on only fast blood glucose without its association with vital signs.

The overall effect of controlling unseen data issues on the hidden one-dimensional data underlying multivariate vital sign data via fusion weights, and statistics based on the choice of similarity measure is crucial in understanding the variability in the operation mechanism of the entire vital signs. This makes the path of an absence statistic for given vital sign to provide a better roadmap for assessing how remaining vital signs can stand in for the entire operation of the ecosystem in the absence of the vital sign in question. This in a way allows a non-trivial approach for checking implication of similarity measures and fusion statistics as well as the fusion method on handling public health data in their ability to offer automatic controls for unseen data issues which may be inevitable. With the above observation, the potential flexibility of the fusion approach to public health problems becomes clearly visible.

References

Charlton, P., Pimentel, M., & Lokhandwala, S. (2016). Data fusion techniques for early warning of clinical deterioration., 325-338. https://doi.org/10.1007/978-3-319-43742-2_22

Dziorny, A., Lindell, R., Fitzgerald, J., & Bonafide, C. (2022). Variations among electronic health record and physiologic streaming vital signs for use in predictive algorithms in pediatric severe sepsis. Aci Open, 06(02), e76-e84. https://doi.org/10.1055/s-0042-1755373

Gao, B., Hu, G., Gao, S., Zhong, Y., & Gu, C. (2018). Multi-sensor optimal data fusion based on the adaptive fading unscented kalman filter. Sensors, 18(2), 488. https://doi.org/10.3390/s18020488

Gao, H., McDonnell, A., Harrison, D. A., Moore, T., Adam, S., Daly, K., ... & Harvey, S. (2007). Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. *Intensive care medicine*, *33*, 667-679. https://DOI 10.1007/s00134-007-0532-3

Habib, C., Makhoul, A., Darazi, R., & Couturier, R. (2017). Real-time sampling rate adaptation based on continuous risk level evaluation in wireless body sensor networks., 1-8. https://doi.org/10.1109/wimob.2017.8115777

Hu, Z., Yang, S., Yang, L., & Jin, Y. (2021). Variational bayesian-based adaptive distributed fusion target tracking with unknown sensor measurement losses. Iet Radar Sonar & Navigation, 16(1), 64-75. https://doi.org/10.1049/rsn2.12164

Jesus, G. and Silva, B. (2022). Robust estimation for discrete-time markovian jump linear systems in a data fusion scenario. Intermaths Revista De Matemática Aplicada E Interdisciplinar, 3(1), 17-36. https://doi.org/10.22481/intermaths.v3i1.10715

Liu, S., Yao, J., & Motani, M. (2019). Early prediction of vital signs using generative boosting via lstm networks., 437-444. https://doi.org/10.1109/bibm47256.2019.8983313

Liu, W., Wang, X., & Deng, Z. (2016). Robust weighted fusion kalman estimators for multisensor systems with multiplicative noises and uncertain-covariances linearly correlated white noises. International Journal of Robust and Nonlinear Control, 27(12), 2019-2052. https://doi.org/10.1002/rnc.3649

Maronna, R. A., Martin, R. D., Yohai, V. J., & Salibián-Barrera, M. (2019). *Robust statistics: theory and methods (with R)*. John Wiley & Sons.

Mensah, D. K., Ofori, M. A., & Howard, N. (2022). Traumatic Systolic Blood Pressure Modeling: A Spectral Gaussian Process Regression Approach with Robust Sample Covariates. *Mathematics and Statisticsi*, 10(3):562–574

Ofori, M. A., Mensah, D. K., Orwa, G. O., & Hewson, P. (2024). Traumatic physiological vital sign fusion: Insight from composite spatial similarity measure modelling. *Authorea Preprints*.

Ofori, M. A., Mensah, D. K., Orwa, G. O., & Hewson, P. (2024). Data recovery methods in composite similarity-based data fusion: Application to Physiological vital signs. *Authorea Preprints*.

Peng, Z., Zhou, S., Liu, P., & Li, M. (2022). Distributed ellipsoidal intersection fusion estimation for multi-sensor complex systems. Sensors, 22(11), 4306. https://doi.org/10.3390/s22114306

Rossum, M., Silva, P., Wang, Y., Kouwenhoven, E., & Hermens, H. (2023). Missing data imputation techniques for wireless continuous vital signs monitoring. Journal of Clinical Monitoring and Computing, 37(5), 1387-1400. https://doi.org/10.1007/s10877-023-00975-w

Sadasivuni, S., Saha, M., Banerjee, I., & Sanyal, A. (2021). Fusion of fully integrated analog machine learning classifier with electronic medical records for real-time prediction of sepsis onset.. https://doi.org/10.21203/rs.3.rs-564329/v1

Scott, D. W. (2015). *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons.

Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.

Skyttberg, N., Vicente, J., Chen, R., Blomqvist, H., & Koch, S. (2016). How to improve vital sign data quality for use in clinical decision support systems? a qualitative study in nine swedish emergency departments. BMC Medical Informatics and Decision Making, 16(1). https://doi.org/10.1186/s12911-016-0305-4

Sun, G., Matsui, T., Watai, Y., Kim, S., Kirimoto, T., Suzuki, S., … & Hakozaki, Y. (2018). Vital-scope: design and evaluation of a smart vital sign monitor for simultaneous measurement of pulse rate, respiratory rate, and body temperature for patient monitoring. Journal of Sensors, 2018, 1-7. https://doi.org/10.1155/2018/4371872

Wang, S., Celebi, M. E., Zhang, Y. D., Yu, X., Lu, S., Yao, X., ... & Tyukin, I. (2021). Advances in data preprocessing for biomedical data fusion: An overview of the methods, challenges, and prospects. *Information Fusion*, *76*, 376-421.

Yan, L., Xia, Y., & Fu, M. (2017). Optimal fusion estimation for stochastic systems with cross-correlated sensor noises. Science China Information Sciences, 60(12). https://doi.org/10.1007/s11432-017-9140-x

Yang, X., Zhang, X., Ding, Y., & Zhang, L. (2021). Indoor activity and vital sign monitoring for moving people with multiple radar data fusion. Remote Sensing, 13(18), 3791. https://doi.org/10.3390/rs13183791

Zhang, Z., Li, J., & Liu, L. (2015). Distributed state estimation and data fusion in wireless sensor networks using multi-level quantized innovation. Science China Information Sciences, 59(2), 1-15. https://doi.org/10.1007/s11432-015-5415-6