Journal of Biomedical Engineering
ISSN: 1001-5515

MANUSCRIPT
Vol. 40 No. 3 (2023)

DATA FUSION FOR PHYSIOLOGICAL VITAL SIGNS WITH AUTOMATIC EXTREME VALUE CONTROL

Francis Eyiah-Bediako[1], David Kwamena Mensah[2], Samuel Assabil[3], Emmanuel Amewonor[4], and Richard Okyere[5]

[1,2,3,4]Department of Statistics, University of Cape Coast, Cape Coast, Ghana

[5]Ghana Insurance College, Accra, Ghana

*Corresponding author: **dmensah@ucc.edu.gh**

Abstract

In this paper, we propose methods for fusing multivariate physiological vital sign data observed over a common time into univariate data. The approaches rely on noncentral moment-based statistics derived from the physiological vital sign random variables. Thus preserve the vital sign specific auto-correlations and appropriately control the within-vital sign extreme observations automatically so that such observations are utilized in the model specification without being deleted. Data mixing factors and weights are developed and allowed to depend on the fusion statistics so as to inherit their appealing features. This way, a single model can be fitted to the composite data instead of multiple models for the unfused data resulting in computational and memory savings. Implementation of the approaches using real vital sign data illustrates the utility of the schemes in extracting appropriate univariate data from the multivariate data that can be handled in terms of the common time within the functional regression framework.

Keywords: Multivariate vital signs, Composite vital sign, Data fusion, Single-task modelling, multi-task modelling, Fusion statistics.

1 Introduction

Physiological vital sign modelling has piqued the interest of many researchers within the statistics community because of its link with health conditions. They have been found to provide insights on early signs of health deterioration, and have been applied for identification of deterioration in health [Gao et al., 2007, Khalid et al., 2013, Pimentel et al., 2013]. Thus, they are considered key elements in health monitoring worldwide [Buist et al., 1999, Evans et al., 2001]. For instance, physiological vital signs are utilized in the assessment of the quality and clinical properties of drugs in public health through empirical statistical measures of overall quantity. A typical example of such statistics is the smoothness index of physiological vital signs used in hypertension management and treatment [Rizzoni et al., 2019, Parati et al., 2014, Mancia and Parati, 2001]. The smoothness index is computed as a ratio between the mean hourly change in vital signs over 24 hours, divided by the standard deviation [Mancia and Parati, 2001]. Interestingly, a practical utility of the above statistic is seen in the assessment and differentiation of antihypertensive agents by the latest hypertension management standard [Mancia and Parati, 2001, Rizzoni et al., 2019]. Furthermore, another line of the application of physiological vital signs is seen in the conventional health monitoring scheme termed the Early Warning (EW) score applied in most Intensive Care

DOI: 10.105515/JBE.40.3.23

Units (ICU) and surgical wards worldwide. Under this scheme, vital signs readings from bedside monitors attached to patients are continuously tracked by nurses with composite numerical scores computed from all the vital signs and applied to prioritize care and treatment [Evans et al., 2001, Gardner-Thorpe et al., 2006].

Alternative to the empirical treatment of health monitoring based on the EW ecosystem is the model-based approaches for which the literature has registered remarkable proposals. Pimentel et al. [2013] identified deterioration in the health of upper-gastrointestinal surgery patients based on their physiological vital signs. Kernel density estimation (KDE) [Silverman, 1986, Scott, 2015] approach was adopted to develop probability density function(pdf)-based monitors for each vital sign characterizing normal and abnormal groups with the grouping conducted using available patient admission information. Inference in terms of decision making was achieved via a comparison of pdf-based novelty scores with a common threshold computed from the normal group. Velardo et al. [2014] successfully prioritized chronic obstructive pulmonary patients for clinical review using kernel density estimation-based vital sign deterioration detectors in the univariate and multivariate dimensions. Khalid et al. [2013] proposed a Gaussian process regression-based deterioration detection in physiological vital signs in the univariate direction. Their approach outperformed an appropriate empirical approach using KDE.

The physiology of humans can be viewed as a complex system consisting of vital organs which work together in a complex manner to define the overall health of an individual. The dynamic inter-relationship existing among these vital organs and health is not fully known or understood by scientists. However, with the knowledge of the characteristic functions of the physiological organs, physicians through convectional ways are able to determine whether a person is sick or otherwise. They do this by examining sample observations from these physiological vital signs presented by patients coupled with other factors such as diet, activity, etc. This provides an empirical conventional way of assessing ailment and treatment. This suggests that the composite of all vital signs is pivotal in making an informed and precise decision on the health condition of a given individual. However, the question of how to derive an appropriate composite vital sign via fusing the individual vital signs together in an unbiased manner whiles preserving the underlying associations existing among these vital signs becomes clearly apparent. Most importantly, for physiological vital signs, autocorrelations can be informative on the underlying vital sign specific dynamics and can be a source of indication of early signs of health-related issues if detected on time. We argue that statistical approaches for fusing multiple vital signs into a univariate vital sign trajectory taking into account the existing autocorrelation among them, while controlling extreme observations stand to improve health monitoring significantly.

Fusing of multi-physiological vital signs generated by systems operating under the same conditions into a univariate vital sign is important in many ways. First, it has the potential to facilitate easy modelling and fitting of statistical models in that the un-fused data may require multiple models as well estimation of unknown parameters, compounding the computational expenses associated with such datasets. Though these multiple models can be fitted simultaneously

DOI: 10.105515/JBE.40.3.23

with the application of some statistical techniques such as multivariate multiple regression etc., the issues with computational overload cannot be ignored. Furthermore, in areas of application where the smoothness index of vital signs and their relationship with available covariates is of interest, such as the management and treatment of hypertension, fused vital signs will provide a principled way to compute composite smoothness index for easy application and parsimony. The combined effect of multiple vital signs can be very informative on some hidden physiological threat or issue that may require urgent attention. This is important during pandemics such as the COVID 19 where pragmatic treatment intervention is of the utmost interest with its effect on human interaction. This paper focuses on lightweight statistical methods for fusing multiple physiological vital signs into a composite univariate vital sign that allows automatic control for extreme observations within vital signs while maintaining the vital sign-specific autocorrelations. To the best of our knowledge, this paper is a premier on the extraction of composite vital sign trajectory from multi-physiological vital signs monitored over common time stamps.

The rest of the paper assumes the following structure. The data fusion model is formally introduced in Section 2. Section 3 introduces in brief, the moment-based approach for selecting appropriate fusing statistics considered in the fusion proposal outlined in Section 2. Section 4 outlines the performance measures for assessing the proposed methods with direction for the implementation of the methods detailed. Real data application using health data is given in Section 5, and Section 7 concludes.

2 Data fusion model

Let $y=\left[y_1,y_2,\ldots,y_p\right]$, $y_j=\left(y_{1j},y_{2j},\ldots,y_{nj}\right)$, $j=1,2,\ldots,p$ denotes a physiological vital signs data associated with a common time stamp. The inter-correlations existing among the defining variables of the data suggests that the variables can be combined appropriately into composite variable. Let y^1 denotes the composite (fused) data generated by y. We consider the following data fusion models for y^1.

$$y_{1i}^t=\theta_i' T\left(y_i\right),\ i=1,2,\ldots,n \qquad (1)$$

$$y_{1i}^v=\theta_i' V\left(y_i\right),\ V\left(y_i\right)=S^2\left(y_i\right) \qquad (2)$$

$$y_{1i}^z=\theta_i' Z\left(y_i\right),\ \mathrm{Z}\left(y_i\right)=\sqrt{V\left(y_i\right)}, \qquad (3)$$

where $T\left(y_i\right)$ are fusion statistics that are to be chosen appropriately with the following notations used.

$$S\left(y_i\right)=\frac{T\left(y_i\right)-u\left(y_i\right)}{\delta_i},\ \delta_i>0$$

$$T\left(y_i\right)=\left[T\left(y_{i1}\right),T\left(y_{i2}\right),\ldots,T\left(y_{ip}\right)\right],\ \mathrm{V}\left(y_i\right)=\left[V\left(y_{i1}\right),V\left(y_{i2}\right),\ldots,V\left(y_{ip}\right)\right]$$

$$T\left(y_i\right)=\left[T\left(y_{i1}\right),T\left(y_{i2}\right),\ldots,T\left(y_{ip}\right)\right],\ f\left(y_i\right)=\left[f\left(y_{i1}\right),f\left(y_{i2}\right),\ldots,f\left(y_{ip}\right)\right]$$

$$\theta_i=\left(\theta_{i1},\theta_{i2},\ldots,\theta_{ip}\right),\ \mathrm{y}_i=\left[y_{i1},y_{i2},\ldots,y_{ip}\right],\delta_i=\left[\delta_{i1},\delta_{i2},\ldots,\delta_{ip}\right]$$

DOI: 10.105515/JBE.40.3.23

The selection of $T(y)$ is discussed in Section 3. The use of the statistics, $V(y_i)$ and $Z(y_i)$ for generating appropriate shape restricted features is motivated by the fact that they have known sampling distributions for large samples. In particular, it can be shown that $Z(y_i)$ and $V(y_i)$ follow standard normal and Chi-square distributions by the Central limit Theorem. In this paper, we consider the estimation of θ_i using $\hat{\theta}_i$

$$\hat{\theta}_i = \frac{T(y_i)}{C^m(y)}, \qquad (4)$$

where $C^m(y) = \left[C^m(y_1), C^m(y_2), \ldots, C^m(y_p)\right]$ is the set of maximum values of $T(y)$. That is, $C^m(y_j) = \max\left[T(y_{1j}), T(y_{2j}), \ldots, T(y_{nj})\right], j = 1, \ldots, p.$ This allows θ_i to be variable specific with the ability to control extreme observations. It is straightforward to see that θ_j in (1), (2) and (3) is not restricted to be a probability due to the use of the standard random variables. One can also consider mixing weights by just converting the θ into probabilities. Mixing weights based on θ of the form ρ can be derived using

$$\rho_i = \frac{|\theta_i|}{\sum_i |\theta_i|}$$

Such that $\sum \rho_j = 1$. Consequently, the corresponding data fusion models based on the mixing weights, ρ_i can be obtained by replacing θ_i with ρ_i in (1), (2) and (3), yielding

$$y_{2i}^t = \rho_i' T(y_i), i = 1, \ldots, n \qquad (5)$$

$$y_{2i}^v = \rho_i' V(y_i) \qquad (6)$$

$$y_{2i}^2 = \rho_i' Z(y_i) \qquad (7)$$

3 Selection of feature *T*(*y*)

The statistic $T(y)$ is crucial in the fusion process since the data fusion models are dependent on it. As a result, its selection must be given a particular attention. Not every statistic can deliver the expected results, since the data exhibits non-linear trends. Also, there may exist extreme observations and inter-relationships among the defining variables of the data. $T(y)$ must be able to capture such information so that the nature of the data is preserved. In particular, the model for the above statistic should be able to handle such features automatically. We consider model (8), for computing $T(y)$.

$$T(y) = yf(y), \qquad (8)$$

$yf(y) = \left[y_1 f(y_1), \ldots, y_p f(y_p)\right]$, with the associated density functions $f(y_j), j = 1, \ldots, p,$ estimated using kernel density estimation. In what follows, we illustrate the kernel density

DOI: 10.105515/JBE.40.3.23

estimation procedure. For y_j with observations $y_{1j}, y_{2j}, \ldots, y_{nj},$, the density function is estimated for each observed value using (9).

$$\hat{f}\left(y_{ij}\right)=\frac{1}{n}\sum_{l=1}^{n}K_{hj}\left(y_{ij}-y_{lj}\right),\ K_{hj}\left(y_{ij}-y_{lj}\right)=\frac{1}{h_j}K\left(\frac{y_{ij}-y_{lj}}{h_j}\right) \qquad (9)$$

for a symmetric kernel function $K_{hj}(.)$ and smoothing parameter, h_j. The kernel density estimation provides a formal approach to estimate the underlying probability density function of a given dataset empirically. For an introduction to kernel density estimation, its implementation, and practical applications, see, for example, [Scott, 2015] and [Silverman, 1986].

The proposal in (8) is motivated by moments of random variables (probability distributions). In particular, for the *j*th random variable Y_j, the *k*th non-central moment is defined as

$$E\left[Y_j^k\right]=\int y_j^k f\left(y_j\right)dy_j \qquad (10)$$

Setting $k=1$, (10) reduces to the usual measure of center termed the expected value. It is not difficult to see that, the integrand, $y_j f\left(y_j\right)$, in (10) provides a principled way to assess the contribution of the observations generated by Y_j. In particular, $f\left(y\right)$ serves as natural weights, penalizing observations according to their numerical strength and spatial positions under the probability density function. This reflects the relative importance of each observation to the common center associated with the density function of Y_j. Practically, the weighting is such that observations in the neighborhood of the common center are weighted less in comparison with those at an appreciable distance from such center. In this regard, it becomes easy to perceive how individual observations are automatically handled within the probability density domain. Thus, outlying or extreme observations will have fair treatment under the probability density function of a given random variable regardless of the nature of the distribution.

Notably, the statistic, $y_j f\left(y_j\right)$ inherits the above vital characteristics of $f\left(y_j\right)$ automatically, making it appealing for generating robust features for data fusion. The presence of outlying observations does not pose any threat to the fusion since they are handled appropriately via $f\left(y_j\right)$. Outlying observations are those with a deviant pattern from the common pattern suggested by the data. They usually appear at the tails of the probability density function that best defines the data. As a result, they have low $f\left(y_j\right)$. values and yield relatively lower contribution values, $y_j f\left(y_j\right)$. Generally, moments of probability distributions or random variables define important features of the distributions of random variables. Thus, the choice of $T\left(y\right)$ based on such statistics, allows $T\left(y\right)$ to inherit the key features of the underlying moment used.

4 Performance assessment

DOI: 10.105515/JBE.40.3.23

We focus on the ability to preserve the underlying pattern, features, and dynamics of the original data in the fused or composite dimension. The fundamental characteristic of an appropriate statistic for extracting a composite data with input being a multi-dimensional data is its order-preserving, pattern preserving capabilities in line with the original data. Of importance, is the pattern preserving characteristics, since any distortion in the composite data may convey different information from that contained in the original data. This may eventually lead to model miss-specification coupled with inappropriate practical implications. The implementation of the methods and algorithms was conducted using the R statistical software. For the probability density-based fusing methods, the kernel density estimation was done using the kernel smoothing package, "ks" in R [Duong et al., 2007]. The smoothing parameter *hi* in (9) was set to the smoothed cross-validation estimator implemented in the above package. All the codes were written in R and run on an Intel (R) Core (TM) *i*7, 6700 processor Windows PC 3.40 GHz workstation.

5 Example: Health data application

We exemplify the utility of the proposed methods with real health data. The dataset is a real physiological vital signs data from Biofourmis, a Health Data Analytic Company. It is multivariate in nature with dimension 848 × 5, generated by variables, Systolic blood pressure (SBP), Diastolic blood pressure (DBP), Mean arterial pressure (MAP), Pulse Rate (PP), and Heart Rate (HR). There exist some patient-specific covariate information on Gender, Age, Height, Weight, and Race, also available. Nevertheless, in this application they were not utilized since they were not useful in the current illustration. The data collection was based on a structured study by SingHeart with a group of Asians. The subjects were continuously monitored over a period based on varied physiological states corresponding to different activities, for example, sleeping, walking, exercising, sitting, etc. The application here utilized a de-identified version of the data in order to address the issue of sensitive patient information. We provide an exploratory analysis of the dataset in brief. The typical nature of each of the corresponding vital signs characterizing the physiological vital sign data are shown in Figure 1. The functional nature of the vital signs trends is evident with each exhibiting varied frequencies according to the intrinsic features of the sample patients. It is clearly obvious from the graphs that some points are substantially deviating from the common functional pattern evident in each of the vital signs data. This departure observed may be due to many hidden reasons or factors, thus, these points can be viewed as potential outliers.

DOI: 10.105515/JBE.40.3.23

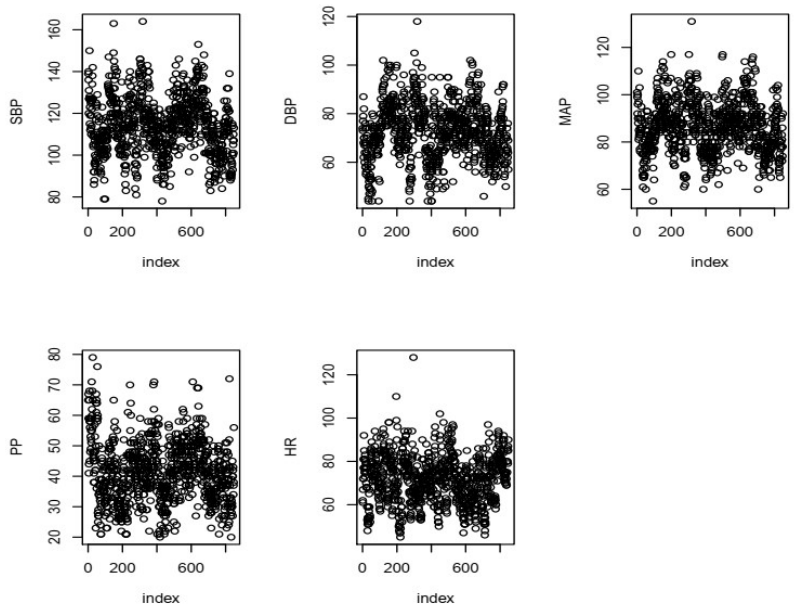

Figure 1: scatter plots of vital signs. From left to right are the plots of SBP, DBP, MAP, PP, and HR respectively.

Figure 2 shows boxplots of vital signs. Clearly, all the vital signs are skewed to the right with Pulse Pressure (PP) exhibiting the highest skewness and Diastolic Blood Pressure being the modest.

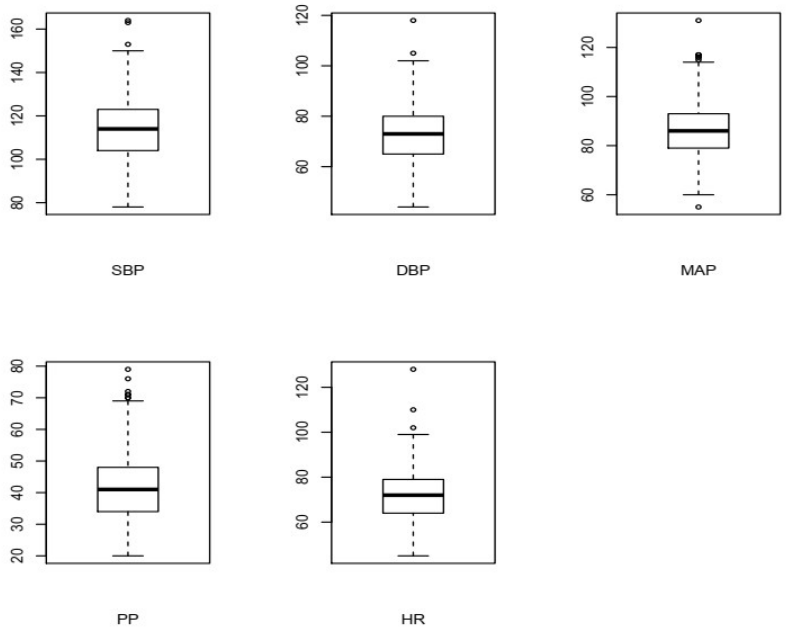

Figure 2: Boxplot of vital signs. From left to right are the plots of SBP, DBP, MAP, PP, and HR respectively.

Figure 3 illustrates the presence of inter-relationships existing among the physiological vital signs. The numbers on the graphs are the Pearson Correlation Coefficients. Clearly, there exist linear

DOI: 10.105515/JBE.40.3.23

relationships among physiological vital signs. For example, SBP is highly positively related to MAP and DBP with Pearson Correlation Coefficient values of 0.82 and 0.70 respectively. It can be observed further that there exists evidence of right skewness of the distributions of the variables. This is evident by the plot of the density functions of the variables as identified in Figure 2.

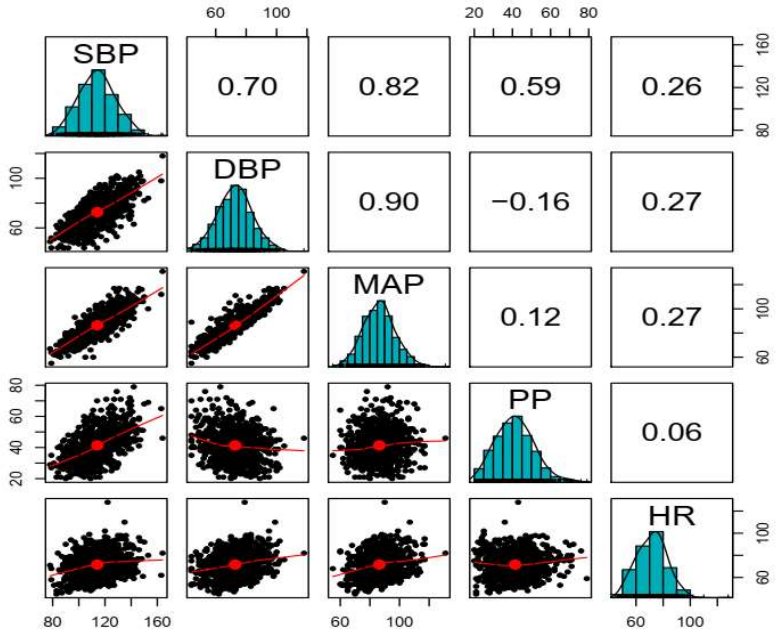

Figure 3: Nature of vital signs data showing density functions, inter-relationships (pearson correlation coefficients).

6 Results and Discussion

The key principle in fusing multivariate functional data into a composite univariate functional data is the ability to obtain common patterns for defining variables of the dataset. As it is well-known in regression analysis that near points offer better prediction of given points than points that are at an appreciable distance. Thus, we argue that similar features from each of the defining variables underlying a p variate multivariate dataset, can ensure appropriate fusing than variable features. Most especially, if the fused pattern is expected to mimic the underlying data structure in order to convey the same information in the original data. For practical purposes, similar features would by no doubt provide a better background for deriving composite patterns, with the ability to allow simple basis for comparison. Figure 4 shows the estimated probability density functions of the standardized (left)and unstandardized (right) vital signs data using model (9). The differences among the vital signs are clearly evident. Also, the value of standardization in obtaining common distributional characteristics is clearly illustrated, evidenced by the level of closeness of the density functions observed in both plots.

Figure 5 shows the nature of the contribution statistic $T(y)$ in relation to each of the vital sign observations for both the standardized (left) and unstandardized (right) data. It can be observed

DOI: 10.105515/JBE.40.3.23

that the $T(y)$ statistic can yield similar (common) pattern for the defining variables of the multivariate dataset with the effect of standardization clearly illustrated. Also, $T(y)$ appears to be a good candidate statistic within the standardized variable context for fusing multivariate vital signs into a univariate vital sign.

Figure 6 illustrates the vital sign-specific contribution pattern underlying the multivariate physiological dataset. The green lines denote the maximum contributions. Clearly, the contribution pattern is nonlinear, adding to the appropriateness of such statistics for use as building block for fusion statistics, since the original data pattern is nonlinear. The differences among the vital sign-specific observations are apparent. Furthermore, it can be observed that the efforts and importance of each observation generated by a vital sign are easily quantified and assessed via the $T(y)$ statistic.

The utility of the $V(y)$ statistic in the extraction of common features across vital signs as a basis for fusing multivariate vital signs into a one-dimensional vital sign is shown in Figure 7. The plot on the left is the plot of $V(y)$ against each of the standardized vital signs. The plot on the right shows the plot of $V(y)$ against each of the unstandardized vital signs. It can be observed that the features generated by this statistic exhibit a similar pattern with the underlying difference existing among the vital signs illustrated high for the unstandardized variables.

Figure 8 shows the underlying patterns of the vital sign-specific $Z(y)$ statistics for both the standardized and unstandardized datasets. Again, clear differences exist among the patterns generated by the two datasets, with the standardized dataset exhibiting more compact patterns than its unstandardized counterpart.

Figure 9 shows the plots of the mixing factors (θ) and weights (ρ) against the vital signs for the standardized dataset. It can be observed that the mixing factors mimic the pattern of the $T(y)$ statistic while the mixing weights inherit the nature of the $V(y)$ statistic. This characteristic of the mixing factors and weights is of key importance in extracting composite physiological vital sign trajectory from a multivariate vital sign dataset where the presence of extreme observations is inevitable. This is because they exhibit the ability to control extreme observations automatically. In addition, the expectation that the output should be functional and mimic the underlying physiological vital signs to enable encoding of the condition of health for pragmatic health monitoring with practical importance. It is important to note that the mixing weights and factors can exhibit variant patterns from those underlying the true vital sign, especially when they are built from sources that do not preserve the order of the observations as well their status in relation to the common data structure underlying the dataset.

Figure 10 shows the plot of the mixing factors and weights against the $T(y)$ statistic based on the standardized dataset. The difference in the relationship existing among the mixing factors and

DOI: 10.105515/JBE.40.3.23

weights and the contribution statistics is evident. Nevertheless, similar patterns are reported by all the vital signs for both the mixing factors and weights.

Figure 11 presents the plot of vital sign-specific highest contributors measured using the $T(y)$ statistic for the standard dataset. Clearly, there exist multiple highest contributors among each of the physiological vital signs. However, they are handled uniquely in such a way that the composite vital sign trajectory is not affected. Most importantly, repeated observations are inevitable in a physiological vital sign dataset and the ability to handle such observations without compromising on their locations and contributions is vital in allowing their intrinsic relationships with the condition of health to be incorporated into modelling. This allows development of intelligent statistical models for providing pragmatic health monitoring schemes. Also, the ability of the proposed methods in handling repeated observations is evident.

Table 1 and Table 2 report the estimates of the smoothing parameters, h_i and variances, δ_i, of the contribution statistic, $T(y_i)$ for the standardized and unstandardized vital sign datasets respectively. Marginal differences exist among the smoothing parameters of the standardized vital signs yielding a similar level of smoothness in the resulting density estimates. On the other hand, significant differences exist among both the $h_i s$ and $\delta_i s$.

Figure 12 shows the vital sign-specific trajectories underlying the physiological dataset. The functional nature of the vital sign dataset is clearly illustrated. Also, the differences in the physiological trajectories are clearly evident.

Figure 13 shows the plot of the composite physiological vital sign trajectory obtained by fusing all the five vital sign trajectories in Figure 12. From left to right of the first row are the plots for y_1^t, y_1^v, and y_1^z respectively. However, row two presents the plots for y_2^t, y_2^v, and y_2^z respectively from left to right. We are able to derive functional univariate physiological vital sign trajectory preserving the common pattern underlying the dataset. Also, the order of the observations is preserved in the composite vital sign trajectory. Nevertheless, there exist some minor differences among the composite trajectory obtained using the mixing factors and weights, evident in the plots in the last two columns. Finally, though the fusing statistics, $T(y)$, $V(y)$ and $Z(y)$ are shape restricted, they do not influence the pattern of the composite vital sign. Thus, the proposed methods preserve the true pattern of that underlies the multivariate physiological vital sign data.

DOI: 10.105515/JBE.40.3.23

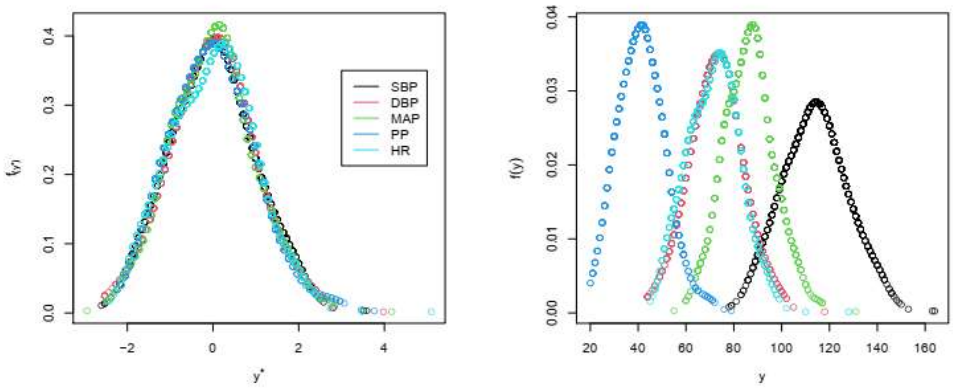

Figure 4: Nature of Vital signs specific density plots for standardized and unstandardized physiological data.

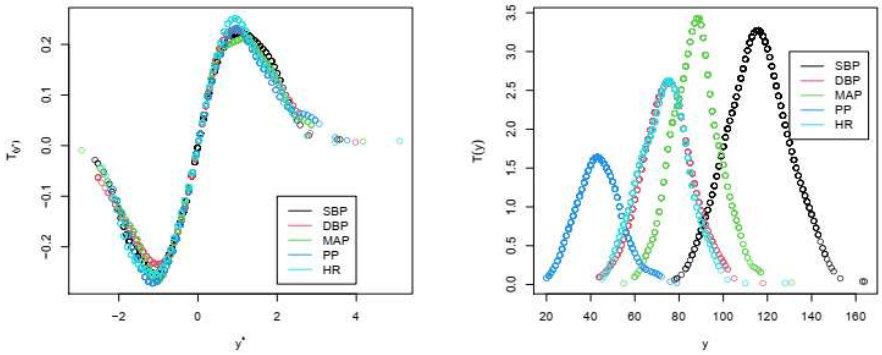

Figure 5: Nature of Vital signs specific $T(y)$ statistics for standardized (left) and unstandardized (right) data.

DOI: 10.105515/JBE.40.3.23

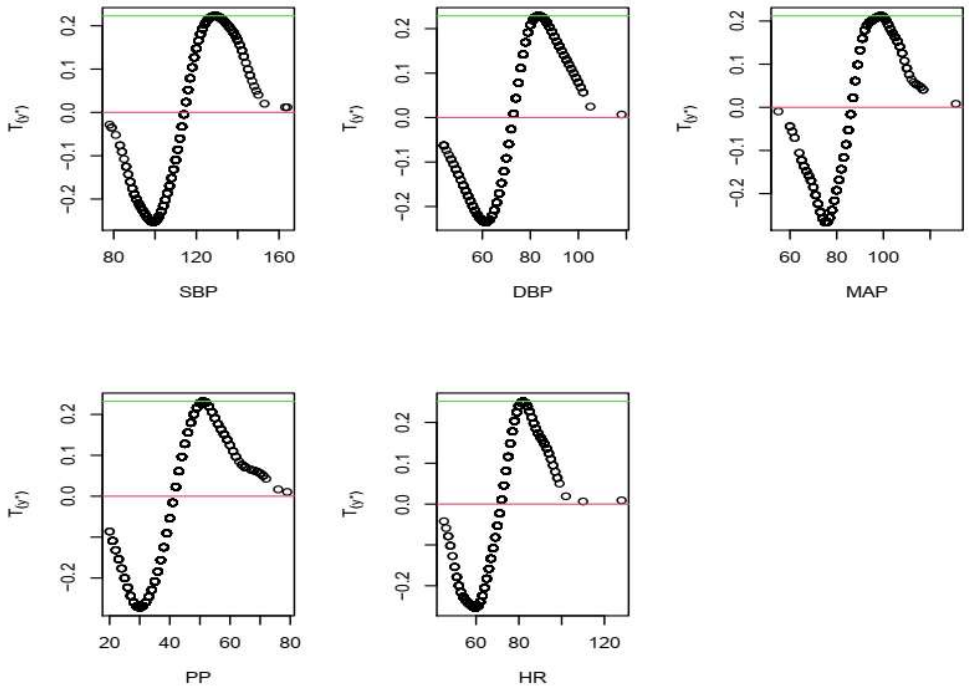

Figure 6: Plot of vital sign specific observation contributions

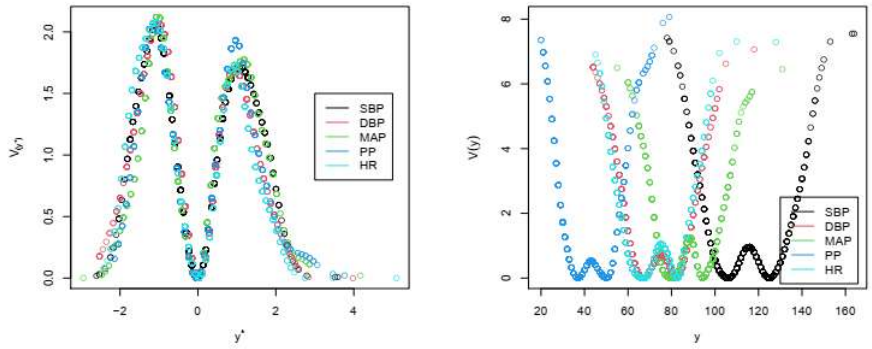

Figure 7: Characteristics of Vital signs specific $V(y^*)$ statistics for standardized and unstandardized physiological data.

DOI: 10.105515/JBE.40.3.23

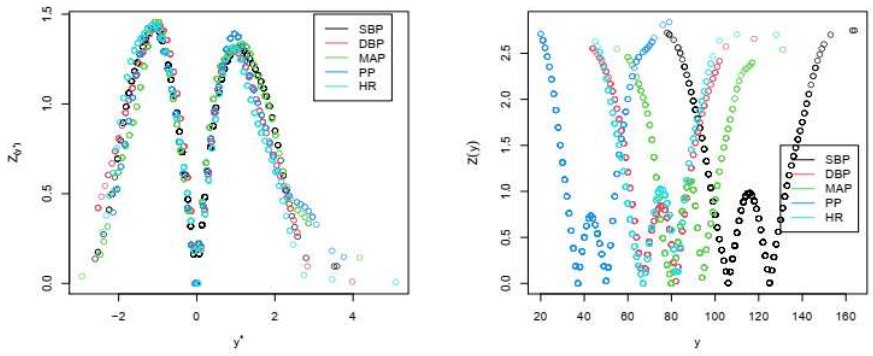

Figure 8: Nature of Vital signs specific $Z(y^*)$ statistics for standardized and unstandardized physiological data.

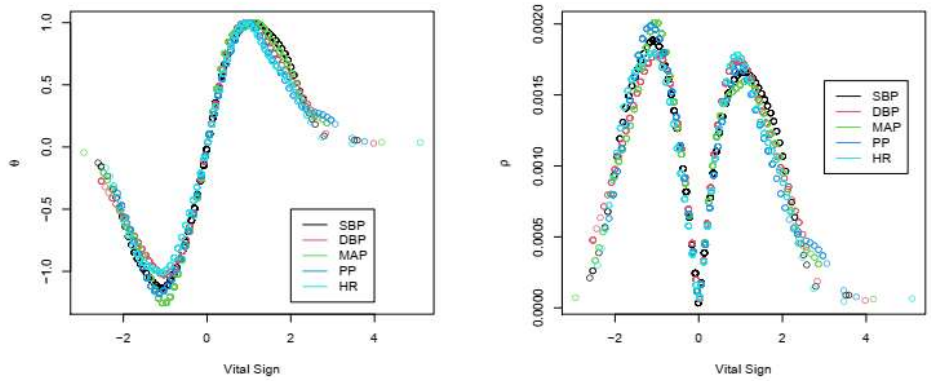

Figure 9: Characteristics of mixing factors and weights, θ, ρ for physiological data.

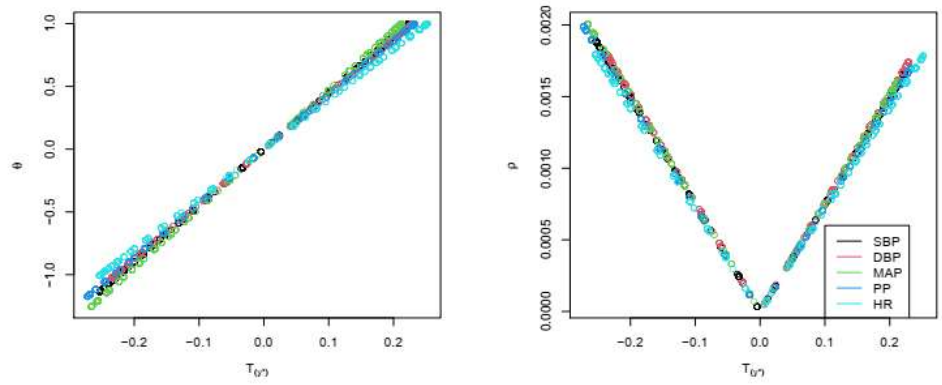

Figure 10: Plot of mixing factors and weights against $T(y^*)$ statistic.

DOI: 10.105515/JBE.40.3.23

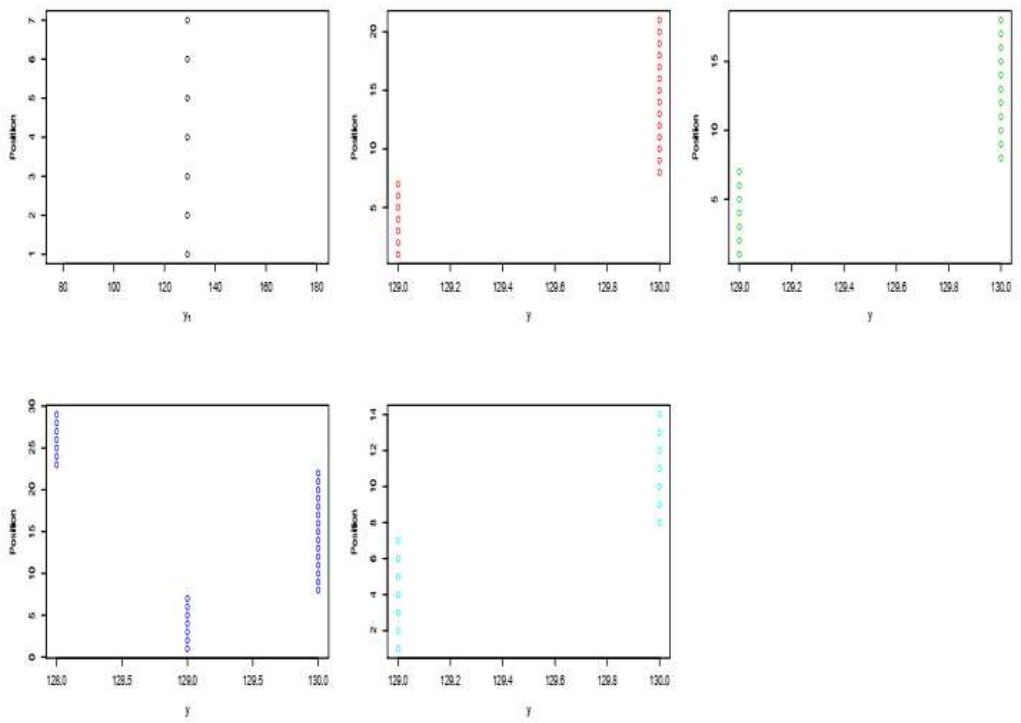

Figure 11: Plot of vital sign observations with maximum contribution, $T(y^*)$ values.

Table 1: Comparison of Vital Sign specific h_i and δ_i for standardized dataset. $y_1^* = \text{SBP}, y_2^* = \text{DBP}, y_3^* = \text{MAP}, y_4^* = \text{PP}, y_5^* = \text{HR}$.

Vital Signs					
	y_1^*	y_2^*	y_3^*	y_4^*	y_5^*
h_i	$\hat{h}_1$	$\hat{h}_2$	$\hat{h}_3$	$\hat{h}_4$	$\hat{h}_5$
	0.2762	0.2731	0.2391	0.2810	0.2606
δ_i	$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{\delta}_3$	$\hat{\delta}_4$	$\hat{\delta}_5$
	0.0301	0.0287	0.0293	0.0318	0.0336

DOI: 10.105515/JBE.40.3.23

Table 2: Comparison of Vital Sign specific h_i and δ_i for unstandardized dataset. $y_1^* = \text{SBP}, y_2^* = \text{DBP}, y_3^* = \text{MAP}, y_4^* = \text{PP}, y_5^* = \text{HR}$.

Vital Signs					
	y_1	y_2	y_3	y_4	y_5
h_i	$\hat{h}_1$	$\hat{h}_2$	$\hat{h}_3$	$\hat{h}_4$	$\hat{h}_5$
	3.8375	3.1075	2.5552	2.8074	2.8798
δ_i	$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{\delta}_3$	$\hat{\delta}_4$	$\hat{\delta}_5$
	0.7522	0.5487	0.8695	0.2061	0.4889

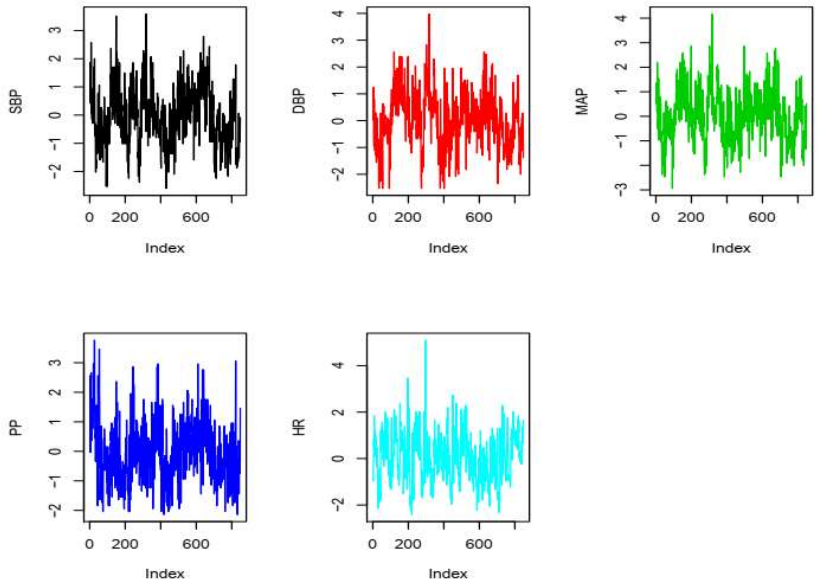

Figure 12: Vital sign specific graph.

DOI: 10.105515/JBE.40.3.23

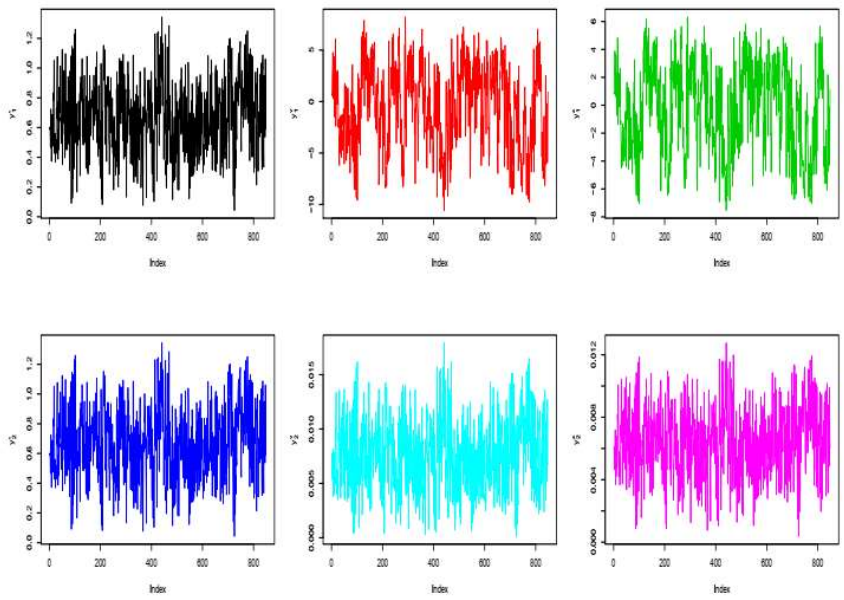

Figure 13: Composite vital signs.

7 Conclusion

This paper proposed and implemented a novel empirical approach for fusing multivariate physiological vital signs data observed over a common time-stamp into univariate physiological vital sign data. The fusing is done through mixing factors, mixing weights, and fusing statistics generated by the vital signs-specific variables such that the within and between vital sign-specific differences in pattern underlying the original data are preserved. The fusing statistics are built using candidates from non-central moments of the random variates such that the probability density function values of observations serve as natural weights, penalizing the observations according to their contributions to the vital sign specific common center associated by the vital signs. The mixing factors and weights are computed as functions of the fusing statistics; thus, allowing them to inherit their desirable intrinsic characteristics.

The probability density functions of vital signs used as a building block of the fusing statistics are estimated using the kernel density estimation approach. This allows the fusing approach to automatically handle vital sign-specific autocorrelation and extreme value control such that all observations are employed without loss of information. The utility of the methods is demonstrated using real multivariate vital signs dataset with common time stamps.

Acknowledgements

The authors acknowledge the support of the Speicial Research Grant RSG/INDI/CANS/2020/131 from the Directorate of DRIC, University of Cape Coast, and Biofourmis, Singapore with real physiological vital signs dataset.

Funding

This research was funded by the Directorate of Research Innovation and Consultancy (DRIC), University of Cape Coast, with the Special Research Support grant RSG/INDI/CANS/2020/131.

DOI: 10.105515/JBE.40.3.23

References

Haiyan Gao, Ann McDonnell, David A Harrison, Tracey Moore, Sheila Adam, Kathleen Daly, Lisa Esmonde, David R Goldhill, Gareth J Parry, Arash Rashidian, et al. Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. *Intensive care medicine*, 33(4):667–679, 2007.

Sara Khalid, David A. Clifton, and Lionel Tarassenko. A bayesian patient-based model for detecting deterioration in vital signs using manual observations. In *International Symposium on Foundations of Health Informatics Engineering and Systems*, pages 146–158. Springer, 2013.

Marco AF Pimentel, David A Clifton, Lei Clifton, Peter J Watkinson, and Lionel Tarassenko. Modelling physiological deterioration in post-operative patient vital-sign data. *Medical & biological engineering & computing*, 51(8):869–877, 2013.

Michael D Buist, Paul R Burton, Stephen A Bernard, Bruce P Waxman, and Jeremy Anderson. Recognising clinical instability in hospital patients before cardiac arrest or unplanned admission to intensive care: A pilot study in a tertiary-care hospital. *Medical Journal of Australia*, 171(1):22–25, 1999.

David Evans, Brent Hodgkinson, and Judith Berry. Vital signs in hospital patients: a systematic review. *International journal of nursing studies*, 38(6):643–650, 2001.

Damiano Rizzoni, Anna Paini, Carolina De Ciuceis, Claudia Agabiti-Rosei, and Massimo Salvetti. The smoothness index: an 'all purposes' approach to the assessment of the homogeneity of 24-h blood pressure control? *Journal of hypertension*, 37(12):2341–2344, 2019.

Gianfranco Parati, Eamon Dolan, Ludwin Ley, and Helmut Schumacher. Impact of antihypertensive combination and monotreatments on blood pressure variability: assessment by old and new indices. data from a large ambulatory blood pressure monitoring database. *Journal of hypertension*, 32(6):1326–1333, 2014.

Giuseppe Mancia and Gianfranco Parati. Importance of smooth and sustained blood pressure control in preventing cardiovascular morbidity and mortality. *Blood pressure*, 10(3):26–32, 2001.

J Gardner-Thorpe, N Love, J Wrightson, S Walsh, and N Keeling. The value of modified early warning score (mews) in surgical in-patients: a prospective observational study. *The Annals of The Royal College of Surgeons of England*, 88(6):571–575, 2006.

Bernard W Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.

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

Carmelo Velardo, Syed Ahmar Shah, Oliver Gibson, Heather Rutter, Andrew Farmer, and Lionel Tarassenko. Automatic generation of personalised alert thresholds for patients with copd. In *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, pages 1990–1994. IEEE, 2014.

 DOI: 10.105515/JBE.40.3.23

Tarn Duong et al. ks: Kernel density estimation and kernel discriminant analysis for multivariate data in r. *Journal of Statistical Software*, 21(7):1–16, 2007.

DOI: 10.105515/JBE.40.3.23