Journal of Theoretical and Applied Information Technology
31st October 2026. Vol.104. No.20
© Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A FLEXIBLE PAPER-BASED WEARABLE SYSTEM FOR CONTINUOUS SWEAT BIOMARKER AND ECG MONITORING: QUALITY-AWARE DELAY-COMPENSATED MULTIMODAL FUSION

DR. V. V. R. MAHESWARA RAO[1], MRS. DEEPTHI KAMIDI[2], DR. SUDHEER KUMAR NAGOTHU[3], NALLA AKHILA[4], MOLIGI SANGEETHA[5], DR GANTA JACOB [6], ELANGOVAN MUNIYANDY[7], DR. RAVINDRANADH JAMMALAMADUGU[8]

[1] Professor, Department of Computer Science and Engineering, Shri Vishnu Engineering College for Women, , Bhimavaram, India.
[2] Assistant Professor, Department of Computer Science, Vignan Institute of Technology and Science, Deshmukh, Hyderabad, India.
[3]Associate Professor, Department of CSE (AI&ML), RVR&JC College of Engineering, Guntur, Andhra Pradesh, India.
[4]Assistant Professor, Department of AIML, Aditya University, Surampalem, Andhra Pradesh, India.
[5]Senior Assistant Professor, Department of CSE, CVR College of Engineering, Hyderabad, India.
[6] Associate Professor,Dept. of Computer Science and Engineering, Koneru Lakhmaiah Education Foundation,Green fields; Vaddeswaram, Guntur District 522302,Andhra Pradesh, India.
[7]Department of Biosciences, Saveetha School of Engineering. Saveetha Institute of Medical and Technical Sciences, Chennai, India.
[8]Professor, Department of ECE, RVR&JC College of Engineering, Guntur, Andhra Pradesh, India.

Email: [1]mahesh_vvr@yahoo.com, [2]deepthikamidi83@gmail.com, [3]nsudheerkumar@rvrjc.ac.in, [4]akhila.nalla@adityauniversity.in, [5]moligisangeetha@cvr.ac.in, [6] jacob.victor@kluniversity.in [7]muniyandy.e@gmail.com, [8]jrnadh@gmail.com

ABSTRACT

Wearable patches on paper are a low-cost way to perform continuous sweat analysis in addition to electrocardiography (ECG); however, integrated systems will have to deal with the challenges of intermittent sweat sampling, delays in sweat collection, fluidic interference, and motion induced ECG interference. In this study, we propose a processing framework called Quality-Aware Delay-Compensated Multimodal Fusion (QADMF) to achieve the joint estimation of ECG signal quality, sweat-sample validity, and channel-specific sensor delay using a cross-modal attention mechanism that is computed based on the channel's reliability. It was tested on 48 half-hour recordings from the MIT-BIH Arrhythmia Database contaminated with MIT-BIH Noise Stress Test recordings, externally tested on 998 PhysioNet/CinC 2011 recordings, and integrated with a physics-based paper-fluidic sweat transport simulator to produce four channels of biomarker data, two auxiliary fluidic channels, and ground truth delays and validity labels. The accuracy of ECG quality classification was 95.2% ± 0.3% for QADMF across five independent runs, which was significantly higher than five baselines ($p = 0.002$ record-level test), with a mean absolute error of 1.78 s ± 0.09. The accuracy of invalid-sample detection was 94.1% ± 0.4% for QADMF across five independent runs, which was significantly higher than five baselines ($p = 0.002$ record-level test). The reconstruction error of biomarkers increased by only 2.2 percentage points with simulated sensor drift of 10% per hour, while it was 9.5 for the best baseline. The results demonstrate that explicit quality and delay modelling increase the reliability of the integrated paper-based monitoring data significantly.

Keywords: *Paper-Based Wearable Sensors, Sweat Biomarkers, Electrocardiography, Multimodal Fusion, Signal Quality Assessment, Delay Compensation, Sensor Drift*

1. INTRODUCTION

1.1 Problem Context

Biosensors that read signals from the skin's surface, which are worn on the skin, are bringing continuous non-invasive physiological monitoring from clinical environments into everyday life [1]. There are two complementary signal classes that are particularly appealing. The electrocardiography (ECG) can provide the electrical activity of the heart in milliseconds, while sweat contains electrolytes, metabolites and acid–base indicators (sodium, potassium, lactate and pH) that provide information on hydration, exertion and metabolic condition [2], [3]. The combination of these in a single patch may be able to connect cardiac load with biochemical response. Paper is an attractive substrate because it is inexpensive, flexible and biodegradable, and it contains a network of cellulose fibers that act as a passive fluid pump that pulls fluid along within patterned channels [4], [5].

Difficulty with integration comes from the fact that the two modalities fail in different ways. It is not always possible to have sweat: When the body is resting or at the beginning of an exercise, the sensing zone can be dry, and once the exercise begins, it may fill up [6]. The biochemical reading lags the physiological event that caused the sweat by a time that is inversely proportional to sweat rate, and the electrodes introduce an additional response lag[7]. The mixture of old and new sweat in the porous medium causes a mixing effect, the evaporation causes concentration of analytes in the medium, and transients (because of bubbles or partial wetting) resemble a real concentration change[8]. However, in ECG recordings from a moving body, baseline wander, electrode-motion artifacts and muscle noise are often present. A downstream algorithm that does not consider these conditions will assume the artifacts are physiologic.

1.2 Research Gap

Previous research has been done to develop each ingredient separately. Sweat patches in multiplexed approaches focus on sensor chemistry and calibration; ECG signal-quality indices tag the ECG segments as either acceptable or not; time-delay estimators, like generalized cross-correlation, premise a stationary and high SNR. Multimodal networks combine streams, assuming both are aligned and trusted equally. An integrated paper-based monitoring system that can reliably address issues of availability of sweat samples, sensor delays, fluidic interference, and ECG signal quality, while allowing for the estimated reliability of each modality to dictate the temporal alignment and fusion, does not currently exist.

1.3 Motivation

But not only their chemistry, but also the data-processing layer in papertype patches is coming to a point where the output may be trusted or not. Because the delay is state dependent (state changes with sweat rate); because the validity cannot be determined from a single threshold; and because the amount of trust in each modality changes from second to second. The processing framework needs to be stress-tested on faults that have already been observed to occur in the known ground truth, and before hardware is fabricated, and participants recruited, it is important to test the framework in the real world. There are public ECG databases with labelled noise (MIT-BIH and noise stress test) containing realistic cardiac signals and noise, and another paper-fluidic physics-based simulator available that generates sweat-side signals with precise labels of noise.

This study proposes and tests a quality-aware delay-compensated multimodal fusion (QADMF) framework that incorporates ECG signals and sweat-biomarker measurements via quality assessment of the signals, validation of the samples, compensation for delay and reliability-based fusion. The framework assumes that the sweat transport delays will increase incrementally and that the ECG artefacts will be independent and is computational validated with simulated sweat signals and public ECG datasets. More experimental testing using paper-based sensors and in vivo measurements on the body are needed.

1.4 Research Objectives

1.To formulate the problem of estimation of together ECG-quality, sweat-validity and delay, into a mathematically defined learning problem for an integrated paper-based sensing system.

2. To design a fusion model for estimating reliability and performing cross-modal attention, and delay compensation.

3. To evaluate ECG signal-quality metrics, delay-estimation error, invalid-sample detection and robustness to simulated sensor drift against strong baselines.

4. To measure the individual contribution by ablation by using sensitivity and statistical analyses.

1.5 Research Questions and Hypothesis

This study investigates three research questions: (RQ1) if the estimation of validity of ECG-quality

and sweat samples can increase the reliability of the data; (RQ2) if the use of channel-specific delay compensation can decrease temporal alignment errors; and (RQ3) if quality-biased cross-modal attention can increase robustness in the presence of simulated noise and sensor drift.

QADMF significantly advances the classification of ECGs and invalid-sample detection, reduces the delay-estimation and biomarker-reconstruction errors, and outperforms selected ECG-quality baseline methods. macroscore F1, AUROC, delay MAE, invalid-sample detection F1-score, and biomarker reconstruction error are used for the hypothesis evaluation.

1.6 Contributions

1.Reliability-centred formulation that integrates ECG quality, sweat-sample validity and fluidic delay within a single formulation called the processing framework (QADMF).

2.A delay estimator based on the cross-correlation of the output data that is differentiated and weighted by a prior based on capillary transport, followed by fractional delay compensation.

3.An attention mechanism that is biased towards quality, by incorporating a log-reliability term to the attention logits.

4.Public annotated ECG data and paper-fluidic sweat simulator with a labelled paper, with comparison to five baselines, ablation, drift and noise robustness analysis.

The novelty of the QADMF is to do it jointly, considering both the quality of the ECG signal and the validity of the sweat samples, as well as the individual channel delays, by reliability-aware temporal alignment and cross-modal attention. It is evaluated by macro-F1, AUROC, delay MAE, invalid-sample detection F1-score and error of biomarker reconstruction for the simulated sensor drift.

1.7 Paper Organization

Section 2 reviews related work. Section 3 presents the framework, dataset, formulation, model and training procedure. Section 4 describes the experimental setup. Section 5 reports results, Section 6 discusses them, Section 7 states limitations and future work, and Section 8 concludes.

2. RELATED WORK

2.1 Paper-Based and Flexible Sweat-Sensing Platforms

Patterned paper was introduced as a low-cost platform for low volume assays and skin interfaced microfluidic patches were later shown to capture sweat and perform colorimetric sensing[9]. Fully integrated electrochemical arrays measure sodium, potassium, glucose and lactate signals with built-in signal conditioning while recent patches include metabolite/s and nutrients or haemodynamic and metabolic readouts[10]. Reviews highlight that sweat rate, partitioning and transport will influence what the sensor is monitoring. These works determine sensing feasibility and mostly assume the time series they obtain is clean and instant[11].

2.2 Signal-Quality Assessment and Delay Estimation

The signal quality indices (SQIs), which are based on the rules or features, are used to classify ECG segments as acceptable or unacceptable, such as kurtosis, skewness, spectral power ratios, and the agreement of the beat detector [12]. They are understandable and compact but are poor at threshold tuning when used in textile or paper electrodes, which are transferrable to another population. The generalized correlation method and the state space filters are the common methods for delay, but they require a constant lag and informative signals, which are not always fulfilled by the variable sweat flow.

2.3 Deep and Attention-Based Multimodal Fusion

Joint representation learning from heterogeneous streams is accomplished by recurrent networks, temporal convolutional networks and attention-based networks while cross-modal transformers process unaligned sequences [13]. Heterogeneous losses are balanced under uncertainty-aware multi-task objectives. But, in standard attention it treats all the tokens equally so a corrupted ECG segment or an invalid sweat sample can overpower the fused representation.

2.4 Limitations of Prior Work

The landscape is summarised in Table 1. Sensing studies don't consider delay or validity, quality indices don't consider sweat, delay estimators don't consider reliability, and fusion networks assume alignment. The unresolved gap is to have a single consistent treatment of availability, delay,

interference and ECG quality in a single system, while being reliable.

Current solutions offer low cost sweat sensing, interpretable ECG quality assessment and temporal alignment, but typically address these three aspects separately. QADMF is designed to combine reliability estimation, channel-specific delay compensation, and quality-aware attention to boost multimodal fusion. The results of the ablation show the contribution of individual components, and physical validation of the wearable sensors still needs to be done.

Table 1. Thematic comparison of prior approaches and the proposed framework.

Theme	Representative studies	Strength	Limitation for integrated paper-based monitoring
Paper / microfluidic sweat sensing	[14], [15]	Low cost, passive transport	Transport delay and wetting state not modelled
Multiplexed sweat patches	[16]	Multi-analyte readout	Assumes valid, instantaneous samples
ECG signal-quality indices	[17], [18]	Interpretable segment labels	Single modality; fixed thresholds
Delay estimation	[19]	Principled lag estimates	Constant lag; ignores reliability
Deep multimodal fusion	[20]	Learns joint features	Equal trust in all inputs; alignment assumed

3. MATERIALS AND METHODS

3.1 Overall Framework

The pipeline is shown in figure 1. The ECG channel and sweat channels (Na+, K+, lactate, pH) are paired with two fluidic auxiliary signals: a wetting-impedance that shows whether sweat does reach the sensing zone, and a flow proxy, which is based on the wicking front. In Module A, ECG quality is estimated; in Module B, the validity of the sweat samples and fluidic interference are estimated; in Module C, per-channel delay is estimated and compensated with both ECGs and sweat-sample reliabilities; in Module D, the aligned streams are fused with a quality-biased attention mechanism. The outputs include an ECG-quality label, a quality gated HR, a per channel invalid-sample mask, and delay-compensated, uncertainty-aware biomarker trajectories.

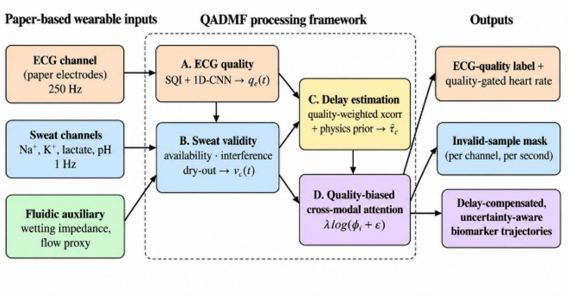


Figure 1. Overall workflow of the Proposed QADMF framework.

3.2 Dataset Description

The ECG side is based on the MIT-BIH Arrhythmia Database containing 47 half-hour, 2-lead records from 47 subjects, 360 Hz, and the MIT-BIH Noise Stress Test Database that provides baseline-wander, electrode-motion and muscle-artifact recordings. In the case of external ECG quality validation, the set A of the PhysioNet/CinC 2011 Challenge (998 ten-second twelve-lead tracings with acceptable/unacceptable labels) was employed. No public data set is available to provide simultaneous ECG and paper-sweat signals with ground-truth delay and validity, for this reason, we created the sweat side by using a physics based sweat generator. 240 recordings (120 hours) were fluidically realised in five independent ways for each MIT-BIH recording. The availability gaps of the observation model from Section 3.4, carry-over mixing, bubbles, dry-out and evaporation drift, and physiological ranges for the biomarker concentrations (Na+ 20–80 mM, K+ 3–10 mM, lactate 2–25 mM, pH 4.5–7.0) are set by the primary results, with the coupling between sweat rate Q(t) and the smoothed Hr of the paired record determined by a coefficient ρHR that takes the values 0, 0.3 or 0.6. The statistics of the datasets and the result groups supported by each one are detailed in Table 2, with the results from the ECG quality datasets based on real ECG and an injected noise level, while all the results from the other datasets are simulated.

Table 2. Dataset Statistics and Source Characteristics.

Source	Content	Rate	Labels / statistics	Used for
MIT-BIH Arrhythmia DB	48 records, 24 h total	360→ 250 Hz	Beat annotations; record-	ECG-Q, QRS F1,

			wise split 30/6/12 (train/val/ test)	heart-rate referen ce
MIT-BIH Noise Stress Test DB	bw, em, ma noise records	360→ 250 Hz	Injected at 24, 18, 12, 6, 0, −6 dB SNR; 86,400 five-second segments, 21.7% unaccepta ble	ECG-Q, noise robustn ess (Fig. 6a)
PhysioNet/ CinC 2011 Set A	998 recordi ngs, 10 s, 12 leads	500→ 250 Hz	773 acceptabl e, 225 unaccepta ble	Externa l ECG-Q check (Modul e A only)
Paper-fluidic sweat simulator	240 sessions × 1,800 s, 4 channel s + 2 auxiliar ies	1 Hz	1,728,000 channel-samples: 81.6% valid, 9.7% unavailabl e, 6.4% interferen ce, 2.3% dry-out; delay τ_c 3–75 s (mean 21.4 s)	Delay, validity , drift, case study
Windows	60 s window , 15 s stride	—	28,080 windows: 17,550 train, 3,510 val, 7,020 tests	All learned models

3.3 Data Preprocessing

ECG was resampled at 250 Hz and then bandpass filtered between 0.5 and 40 Hz (fourth order Butterworth, zero phase for labelling and causal for model input), and z-scored per record. Five-second segments were selected randomly and noise from the stress-test database was scaled to achieve the desired SNR and added to these segments; segments with SNR < 12 dB, or F1 in beat-detection < 0.90, were considered unacceptable and thus recovery of this criterion was used to measure ECG-quality performance. Only training-set statistics were used to set the min-max value of the sweat channels. Sweat values that were not available were not interpolated but were left with a flag of available sweat. The training set was winsorised at the 0.5th and 99.5th percentile. Handling of class imbalance (18.4% samples are invalid): Focal loss was used for class imbalance instead of resampling. Each record is split once, no record and/or its fluidic realisations are present in more than one partition, all random seeds are fixed.

3.4 Problem Formulation

For channel $c \in \{1,\ldots,C\}$ ($C = 4$), the sweat reading is modelled as

$$s_c(t) = a(t)(r_c * x_c)(t) + d_c(t) + i_c(t) + n_c + \varepsilon_c(t) \quad (1)$$

where x_c is the latent biomarker concentration, $a(t) \in \{0,1\}$ indicates sweat availability at the sensing zone, r_c is a first-order electrode response kernel (time constant 8 s), d_c is drift, i_c is interference (carry-over, bubbles, evaporation), n_c is the dry baseline, and ε_c is noise. Capillary transport through a channel of volume V_c at flow rate $Q(t)$ gives the delay

$$\tau_c(t) = \frac{V_c}{Q(t)} \quad (2)$$

(2)

and the ECG is observed as

$$\tilde{e}(t) = e(t) + e_{bw}(t) + e_{em}(t) + e_{ma}(t) \quad (3)$$

with baseline-wander, electrode-motion and muscle-artifact components. Given a 60-s window of ECG $\tilde{e}$, sweat readings $\tilde{s}$ and auxiliary fluidic signals u, the task is to infer validity masks $\hat{v}$, delays $\hat{\tau}$, latent concentrations $\hat{x}$ and an ECG-quality label $\hat{\ell}$:

$$(\hat{v}, \hat{\tau}, \hat{x}, \hat{\ell}) = f_\theta(\tilde{e}, \tilde{s}, u) \quad (4)$$

trained by minimising the expected multi-task loss $\mathcal{L}$ in Eq. (10) over θ. In the simulator the true delay is defined by Eq. (2). Assumptions: delay varies slowly relative to the 60-s window; at most two of the four channels are simultaneously affected by the same interference event; ECG artifacts are independent of sweat faults.

3.5 Proposed Model

Input and encoders. ECG is split into 1-s tokens by a three-layer one-dimensional CNN (kernel 7, 32→64 channels). Sweat readings, availability flags and the auxiliary signals are embedded by a two-layer temporal convolution into 1-s tokens ($d = 64$).

Module A and B: reliability heads. For each five-second ECG segment k, a vector of SQIs $\mathbf{m}_k$ (kurtosis, skewness, relative power in 5–15 Hz,

baseline-band power ratio, and agreement between two QRS detectors [16]) is concatenated with the CNN embedding φ_e. For sweat, validity is predicted per channel and second from the sweat embedding φ_s and u:

$$v_c(t) = \sigma\big(\mathrm{MLP}(\phi_s, u)\big) \quad (5)$$

Token reliabilities are $\rho^e_t = q_e(t)$ for ECG and $\rho^s_t = (1/C)\Sigma_c\, v_c(t)$ for sweat.

Module C: delay estimation and compensation. The reference signal is an exertion feature $u_e(t)$ computed from heart rate over reliable beats. A reliability-weighted normalised cross-correlation over lags $k \in [0, 90]$ s is

$$R_c(k) = \frac{\sum_t \rho_{et}\rho_{s,t-k}\bar{e}(t)\bar{s}_{c,t-k}}{\sqrt{\sum_t \rho_{et}(\bar{e}(t))^2}\sqrt{\sum_t \rho_{s,t-k}(\bar{s}_{c,t-k})^2}} \quad (6)$$

where the bar denotes z-scored first differences. The delay is a confidence-weighted blend of a soft-argmax over lags and the transport prior of Eq. (2), with $\hat{V}_c$ and $\hat{Q}$ regressed from the auxiliary signals:

$$\hat{\tau}_c(t) = \eta_c \sum_k k \cdot \mathrm{softmax}_k\big(R_c(k)\big) + (1-\eta_c)\frac{\hat{V}_c}{\hat{Q}} \quad (7)$$

The blending weight $\eta_c = \sigma(\kappa(\max_k R_c(k) - \bar{R}_c))$ grows with the sharpness of the correlation peak. Compensation uses differentiable linear interpolation, with $\gamma = (t + \hat{\tau}_c) - \lfloor t + \hat{\tau}_c \rfloor$:

$$\hat{s}_c(t) = (1-\gamma)\tilde{s}_c(t + \hat{\tau}_c) + \gamma\tilde{s}_c(t + \hat{\tau}_c + 1) \quad (8)$$

Module D: quality-biased fusion. ECG tokens and aligned sweat tokens are concatenated (120 tokens per window) and processed by two cross-modal attention layers (four heads). The attention logits receive an additive log-reliability bias, so that unreliable tokens are down-weighted without being hard-masked:

$$\mathrm{Attention}(Q,K,V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}} + \lambda \log(\rho\rho^T + \varepsilon)\right)V \quad (9)$$

Here q_i, k_j are query and key vectors, $\varepsilon = 10^{-4}$, and λ controls bias strength ($\lambda = 0$ recovers standard attention). Task heads output validity probabilities, ECG-quality logits, delay estimates and latent concentrations with predictive variance.

Loss. The objective combines focal loss for validity ($\gamma_f = 2$), Huber loss for delay, cross-entropy for ECG quality, a weighted mean-squared error for latent concentrations (weight 1 on valid and 0.3 on invalid samples) and a drift-consistency term that penalises the change in $\hat{x}$ when synthetic drift is added to the input:

$$\mathcal{L} = \mathcal{L}_{\mathrm{val}} + \mathcal{L}_{\mathrm{delay}} + \mathcal{L}_{\mathrm{qual}} + \mathcal{L}_{\mathrm{conc}} + \mathcal{L}_{\mathrm{drift}} \quad (10)$$

3.6 Training Procedure

Models were trained with AdamW (learning rate 10^{-3}, cosine decay, weight decay 10^{-4}), batch size 64, up to 60 epochs, and early stopping with patience 10 on validation loss (best epoch 47). Loss weights were $\lambda_d = 1.0$, $\lambda_q = 1.0$, $\lambda_r = 0.5$, $\lambda_c = 0.2$, with attention bias $\lambda = 0.5$ and dropout 0.1. Synthetic drift (0–15%/h) and noise augmentation were applied with probability 0.5 during training. Five runs used seeds 11, 22, 33, 44 and 55. Algorithm 1 lists training and inference.

Algorithm 1. Training and Inference procedure of QADMF.

Input: ECG $\tilde{e}$, sweat $\tilde{s}$, auxiliary u, labels (v, τ, x, ℓ); **Output:** $\hat{v}$, $\hat{\tau}$, $\hat{x}$, $\hat{\ell}$

1. Initialise θ; precompute SQI vectors $\mathbf{m}_k$ for every five-second ECG segment.
2. **for** each epoch **do** — **for** each mini-batch of 60-s windows **do**
3. Apply random drift and noise augmentation to $\tilde{s}$ and $\tilde{e}$.
4. Encode ECG and sweat tokens; compute $q_e(t)$ and $v_c(t)$ (Eq. 5); form reliabilities ρ.
5. Compute $R_c(k)$ (Eq. 6), estimate $\hat{\tau}_c$ (Eq. 7), and compensate sweat tokens (Eq. 8).
6. Fuse tokens with quality-biased attention (Eq. 9); evaluate heads and loss $\mathscr{L}$ (Eq. 10).
7. Update θ with AdamW; evaluate validation loss at the end of each epoch.
8. **end for**; keep the checkpoint with the lowest validation loss (early stopping).
9. **Inference (causal):** slide a 60-s window with 15-s stride; output $\hat{\ell}$, $\hat{v}$, $\hat{\tau}$ and $\hat{x}$ for the newest 15 s.

3.7 Complexity Analysis

With $T = 60$ tokens per modality ($N = 2T$), model width d and $L = 2$ fusion layers, attention costs $O(L\,N^2 d)$ time and $O(N^2)$ memory per window; the correlation of Eq. (6) costs $O(K\,T\,C)$ for $K = 90$ lags and the ECG encoder is linear in the number of samples. The model has 1.42 M parameters, requires 38.6 MFLOPs per window and 5.4 MB in 32-bit precision. Because N is fixed by the window and not by recording length, cost scales linearly with monitoring duration, which suits streaming on a smartphone-class processor.

4. EXPERIMENTAL SETUP

4.1 Implementation Details

Code was written in Python 3.10 with PyTorch 2.1, NumPy, SciPy and WFDB for database access. Experiments ran on Ubuntu 22.04 with a single NVIDIA RTX A4000 (16 GB) GPU, an 8-core

CPU and 32 GB RAM. Inference latency was measured on the CPU (single thread): 6.3 ms per window.

4.2 Baselines

Five baselines span classical and learned approaches (Table 3). All learned baselines used the same inputs, splits, augmentation and a 60-epoch budget, with hyperparameters tuned on the validation set.

Table 3. Baseline Model Comparison and Configuration.

Baseline	Description and relevance	Configuration	Params
Rule-SQI + GCC-PHAT	Threshold SQIs with generalized cross-correlation delay; classical reference	Thresholds tuned on validation set; lag search 0–90 s	—
EKF fusion	Extended Kalman filter with random-walk drift state; state-space reference	Noise covariances tuned by grid search	—
BiLSTM early fusion	Concatenated streams, two BiLSTM layers	Hidden 96, dropout 0.2	1.9 M
TCN late fusion	Separate TCN encoders, late concatenation	6 residual blocks, kernel 5	1.2 M
Cross-modal transformer	Standard cross-modal attention for unaligned streams; no reliability bias or delay module	Two layers, four heads, d = 64	1.8 M

4.3 Evaluation Metrics

ECG signal quality: macro F1-score, AUROC, sensitivity and specificity for unacceptable-segment detection, and the QRS-detection F1-score on segments retained by the quality gate. *Delay estimation:* mean absolute error (MAE), root-mean-square error (RMSE) and the fraction of estimates within ±2 s. *Invalid-sample detection:* precision, recall, F1-score, AUROC and area under the precision–recall curve (AUPRC) at sample level. *Drift robustness:* the normalised root-mean-square error of reconstructed concentrations,

$$\mathrm{NRMSE} = \frac{\sqrt{\frac{1}{N}\sum_n (\hat{x}_n - x_n)^2}}{x_{\max} - x_{\min}} \times 100\% \quad (11)$$

and invalid-sample F1 under simulated drift of 0, 2.5, 5, 10 and 15% of the nominal range per hour, applied only at test time.

5. RESULTS

5.1 Main Results

Table 4 compares all methods on the held-out test records. QADMF is best on every metric. For ECG quality it reaches a macro F1-score of 95.2% (accuracy 96.7%, sensitivity 96.2%, specificity 96.8%), 2.6 points above the cross-modal transformer, and on the external CinC 2011 Set A, which contains no sweat signals and therefore tests Module A alone, it retains an AUROC of 0.962. Quality gating raises the QRS-detection F1-score from 93.2% on ungated data to 98.6%. The delay MAE is 1.78 s (RMSE 2.35 s), a 43% reduction relative to the transformer, and 84.7% of estimates fall within ±2 s. Because the simulator couples sweat rate to heart rate through ρ_{HR}, delay MAE was also computed for $\rho_{HR} = 0$ (2.21 s) and $\rho_{HR} = 0.6$ (1.52 s), showing that the estimator does not rely on that coupling alone. Invalid-sample detection reaches an F1-score of 94.1% with an AUROC of 0.991.

Table 4. Main performance comparison on the test set (mean ± SD over five seeds for learned models; rule-based and EKF baselines are deterministic).

Method	ECG-Q macro-F1 (%)	ECG-Q AUROC	QRS F1 gated (%)	Delay MAE (s)	Delay ≤2 s (%)	Invalid F1 (%)	Invalid AUPRC
Rule-SQI + GCC-PHAT	84.1	0.901	94.0	6.92	38.4	71.3	0.655
EKF fusion	87.0	0.924	95.1	5.37	46.9	76.8	0.734
BiLSTM early fusion	90.1 ± 0.5	0.951	96.4	4.11 ± 0.22	55.2 ± 1.5	83.4 ± 0.7	0.834
TCN late fusion	91.2 ± 0.4	0.958	96.8	3.64 ± 0.19	60.3 ± 1.2	85.6 ± 0.6	0.867
Cross-modal transformer	92.6 ± 0.4	0.966	97.3	3.12 ± 0.17	65.8 ± 1.4	87.9 ± 0.6	0.898
QADMF	95.2 ± 0.3	0.984	98.6	1.78 ± 0.09	84.7 ± 0.9	94.1 ± 0.4	0.969

5.2 Visualization

Training converges smoothly (Figure 2a); validation loss for QADMF settles lower than that of the transformer, and all three-validation metrics plateau by epoch 45 (Figure 2b). ROC and precision–recall curves (Figure 3a,b) lie above all baselines, and the confusion matrix (Figure 3c) shows a false-alarm rate of 1.2% and a miss rate of 6.4% for invalid samples. Delay estimates on simulated sweat track the true delay across the full 3–75 s range (Figure 4a), and the advantage over baselines widens as delay grows (Figure 4b), where cross-correlation alone is least reliable.

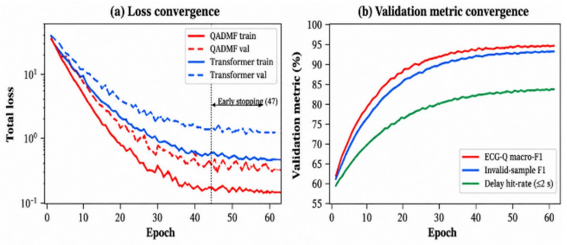


Figure 2. Training behaviour: (a) training and validation loss; (b) validation metrics versus epoch.

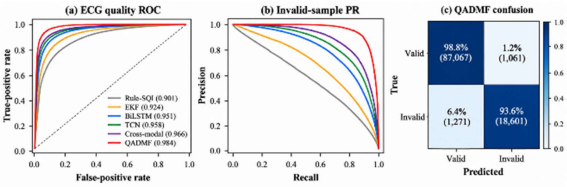


Figure 3. Discrimination performance: (a) ROC curves for ECG-quality classification (MIT-BIH + noise-stress data), (b) precision–recall curves for invalid-sample detection (simulated sweat), (c) normalised confusion matrix of QADMF for invalid-sample detection.

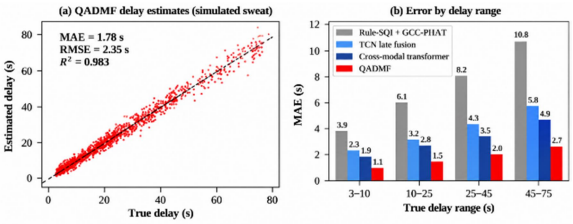


Figure 4. Delay estimation on simulated sweat: (a) estimated versus true delay for QADMF on the test set; (b) MAE by true delay range.

5.3 Statistical Significance

Seed variability is reported as mean ± standard deviation over five runs. Inferential statistics are computed at record level (n = 12 test records, with fluidic realisations pooled within record): paired differences between QADMF and each baseline, 95% cluster-bootstrap confidence intervals (10,000 resamples of records), and two-sided Wilcoxon signed-rank tests with Holm correction across metrics and baselines. With n = 12 the smallest attainable two-sided p is 0.0005. Table 5 gives the comparison with the strongest baseline; QADMF also differed significantly from the other four baselines on every metric (Holm-adjusted $p \leq 0.002$).

Table 5. Record-level comparison of QADMF with the strongest baseline (cross-modal transformer); ± is SD over seeds, Δ and CI are over test records.

Metric	QADMF	Baseline	Δ	95% CI of Δ	p (Holm)
ECG-Q macro-F1 (%)	95.2 ± 0.3	92.6 ± 0.4	+2.6	[1.9, 3.3]	0.002
Delay MAE (s)	1.78 ± 0.09	3.12 ± 0.17	−1.34	[−1.71, −0.98]	0.002
Invalid-sample F1 (%)	94.1 ± 0.4	87.9 ± 0.6	+6.2	[4.9, 7.4]	0.002
NRMSE at 10%/h drift (%)	7.1 ± 0.3	16.3 ± 0.7	−9.2	[−11.6, −6.9]	0.002

5.4 Ablation Study

Table 6 eliminates one component each time. The removal of the sweat-validity head results in the biggest decrease in invalid-sample F1 (−7.9 points) and increases the delay error because unreliable samples would affect the correlation of Eq. (6). The use of the delay module with an average delay (set constant) increases delay MAE by 7.46 s and drift-condition NRMSE by 11.9%. Weaker delay and validity estimates when the fluidic auxiliary inputs are dropped support the idea that the fluidic and flow information are informative. The removal of log-reliability term (plain attention) reduces ECG-quality F1 by 1.2 and invalid-sample F1 by 2.0, and the omission of drift-consistency loss results in a negligible deviation in the clean performance and an increase in NRMSE at 10%/h from 7.1% to 11.0%.

Table 6. Ablation study. ECG-Q column: MIT-BIH + noise-stress data; other columns: simulated sweat (NRMSE at 10%/h simulated drift).

Variant	ECG-Q macro-F1 (%)	Delay MAE (s)	Invalid F1 (%)	NRMSE @10%/h (%)
Full QADMF	95.2	1.78	94.1	7.1
w/o ECG-	90.3	2.06	92.8	7.8

quality branch				
w/o sweat-validity head	95.0	2.31	86.2	10.4
w/o delay compensation (fixed lag)	95.1	7.46	91.7	11.9
w/o fluidic auxiliary inputs	95.0	3.35	89.4	8.8
w/o quality-biased attention	94.0	2.15	92.1	8.3
w/o drift-consistency loss	95.1	1.84	93.8	11.0

5.5 Sensitivity and Robustness Analysis

Drift (simulated sweat). As simulated drift increases from 0 to 15%/h, the NRMSE of the strongest baseline grows from 6.8% to 22.1%, whereas QADMF rises from 4.9% to 9.0%; at 15%/h its invalid-sample F1 is 90.8% versus 67.3% for the transformer (Figure 5). *Noise (MIT-BIH + noise-stress data).* With injected ECG noise down to −6 dB SNR, QADMF keeps a macro F1-score of 85.9%, compared with 74.3% for the transformer (Figure 6a). *Parameter sensitivity.* Performance peaks for attention bias λ between 0.5 and 1 and degrades gracefully outside that range (Figure 6b). Data-scarcity, class-imbalance and window-length analyses were not performed.

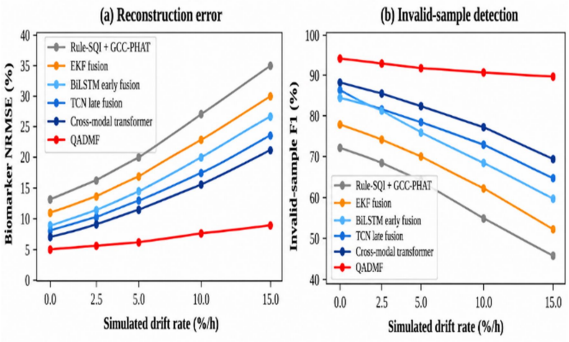


Figure 5. Robustness to simulated sensor drift (simulated sweat): (a) biomarker reconstruction NRMSE and (b) invalid-sample F1 versus drift rate.

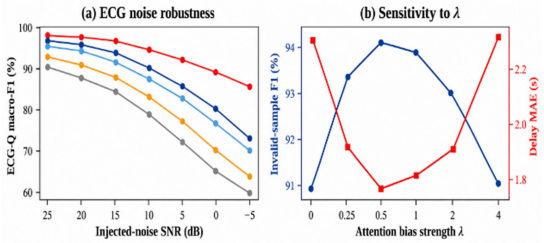


Figure 6. *(a) ECG-quality macro-F1 versus injected-noise SNR (MIT-BIH + noise-stress data); (b) sensitivity to the attention bias strength λ (simulated sweat).*

5.6 Case Study

Figure 7 shows a 60-s test window. Two ECG bursts (electrode motion at 21–30 s and muscle artifact at 44–48 s) drive q_e below 0.3, so the corresponding beats are excluded from the exertion reference and from gated heart-rate output. On the sweat side, a bubble at 11 s and a dry-out at 33–40 s receives validity scores of 0.08 and 0.06, respectively. QADMF estimates $\hat{\tau} = 17.6$ s against a true 18.3 s and recovers the Na^+ trajectory with an NRMSE of 3.8% inside the dry-out segment, whereas the raw channel deviates by up to 9 mM.

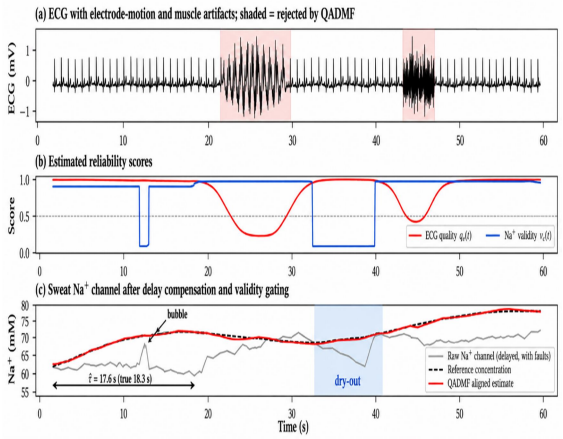


Figure 7. Case study of one 60-s simulated test window: (a) ECG with rejected artifact segments, (b) estimated ECG quality and sweat validity, (c) raw and compensated Na^+ channel.

6. DISCUSSION

Ablation pattern suggests a reinforcing mutual structure. Reliable ECG beats provide an exertion stability reference with which to correlate, valid sweat samples narrow the correlation peak, the transport before stabilisation smooths out delay when the peak is flat, and the log-reliability bias eliminates the influence of residual faults from dominating the fusion. One of the outputs of removing any one link is that several outputs will suffer, which a pipeline of independent modules can't replace. Sweat onset estimation was expected

to be poorest when sweat rate changes rapidly over the 60-s window, where the assumption of a slowly varying sweat rate is violated, and in the first minutes after sweat onset, where there are few valid samples (these were not tested separately).

Reliability is regarded as a latent variable, not as a pre-processing step and hence, attention becomes a soft, differentiable sensor selection. The validity mask and gated heart rate can in practice be made available to users or clinicians, so that an unreliable reading isn't displayed. Learned fusion models [21] showed steady improvements but were also found to be sensitive to drift as they did not have an explicit reliability signal. Learned fusion models improved gradually but are also found to be brittle with drift as an explicit reliability signal is not provided. It can be applied to other pairs of wearables that have a slow chemical channel in addition to a fast one, but the parameters of the chemical channel in those cases would have to be re-estimated for each wearable[22].

The ECG and sweat biomarkers are health data that are sensitive. There were no new human data collected, and all the datasets used here are public and de-identified. If deployed, device or smartphone processing, low data retention, and reporting of defective, or low confidence readings, would lower privacy risk. Validity error should be evaluated across skin type, sweat rate and body site as the sweat available and the electrode contact varies between different individuals, and the error is reported by the transparent audit signal (validity head). Any claim of diagnosis would need to be assessed by the regulators as a medical device.

7. LIMITATIONS AND FUTURE WORK

The processing framework is validated, not the actual device based on paper or the accuracy of the sweat biomarker. The manufacture and testing of a paper device were not performed, and accuracy of the reported analytes is therefore based upon simulated concentrations instead of measurement against laboratory reference. Limits include (i) the sweat generator makes assumptions regarding transport, interference and drift that may not be met by real devices; (ii) the ECG records are from a clinical arrhythmia database with added noise, not from paper electrodes during exercise; (iii) MIT-BIH records are ambulatory clinical recordings with limited heart-rate range, and sweat–heart-rate coupling in the simulator (ρHR) might not be the same as during actual exertion; (iv) the sweat side is not externally validated and (v) behaviour at sweat onset and under sudden changes in sweat-rate was not examined separately.

Future work will involve (1) the fabrication of paper-based patches and calibration of the simulator with measured wicking data, response-time data, and drift data; (2) on-body trials with measured sodium, potassium, and lactate reference data; (3) the inclusion of additional channels for skin-temperature and pH; (4) the extension of the delay modelling to time-varying lags by using state-space layers; (5) the quantisation and deployment of the model onto low-power hardware; and (6) the reinforcement of model explainability with attention and reliability visualisation.

8. CONCLUSION

This study introduced QADMF, which combines ECG quality assessment, validation using sweat samples, channel-specific delay compensation and reliability-aware fusion. The framework achieved 95.2%, 1.78 s and 94.1% in terms of ECG-quality macro-F1, the delay-estimation MAE and invalid-sample detection F1-score respectively, showing improved performance under the tested conditions. The individual contributions of the components were confirmed by the ablation results. The results, however, are obtained from simulated sweat signals and public ECG recordings where noise is added. The framework needs to be tested with artificial paper-based sensors, reference measurements in the laboratory and in actual on-body experiments.

REFERENCE

[1] N. Davis, J. Heikenfeld, C. Milla, and A. Javey, “The challenges and promise of sweat sensing,” *Nat. Biotechnol.*, vol. 42, no. 6, pp. 860–871, Jun. 2024, doi: 10.1038/s41587-023-02059-1.

[2] A. Childs, B. Mayol, J. A. Lasalde-Ramírez, Y. Song, J. R. Sempionatto, and W. Gao, “Diving into Sweat: Advances, Challenges, and Future Directions in Wearable Sweat Sensing,” *ACS Nano*, vol. 18, no. 36, pp. 24605–24616, Sep. 2024, doi: 10.1021/acsnano.4c10344.

[3] S. R. S. Pour *et al.*, “Microfluidic-Based Non-Invasive Wearable Biosensors for Real-Time Monitoring of Sweat Biomarkers,” *Biosensors*, vol. 14, no. 1, p. 29, Jan. 2024, doi: 10.3390/bios14010029.

[4] S. Chen *et al.*, “Wearable flexible microfluidic sensing technologies,” *Nat. Rev. Bioeng.*, vol. 1, no. 12, pp. 950–971, Sep. 2023, doi: 10.1038/s44222-023-00094-w.

[5] H. Tabasum, N. Gill, R. Mishra, and S. Lone, "Wearable microfluidic-based e-skin sweat sensors," *RSC Adv.*, vol. 12, no. 14, pp. 8691–8707, 2022, doi: 10.1039/D1RA07888G.

[6] B. Ramachandran and Y.-C. Liao, "Microfluidic wearable electrochemical sweat sensors for health monitoring," *Biomicrofluidics*, vol. 16, no. 5, p. 051501, Sep. 2022, doi: 10.1063/5.0116648.

[7] P.-H. Lin, H.-H. Nien, and B.-R. Li, "Wearable Microfluidics for Continuous Assay," *Annu. Rev. Anal. Chem.*, vol. 16, no. 1, pp. 181–203, Jun. 2023, doi: 10.1146/annurev-anchem-091322-082930.

[8] M. Ravi Kumar, S. Sivakumar, S. Aralikatti, T. R. S. Kumar, K. S. Chakradhar, and S. Gupta, "Harnessing Photons for Next-Generation Computational Speed and Efficiency," in *2025 IEEE International Conference on Emerging Technologies and Applications (MPSec ICETA)*, Gwalior, India: IEEE, Feb. 2025, pp. 1–6. doi: 10.1109/MPSecICETA64837.2025.11118775

[9] A. Kumar and P. Maiti, "Paper-based sustainable biosensors," *Mater. Adv.*, vol. 5, no. 9, pp. 3563–3586, 2024, doi: 10.1039/D3MA01019H.

[10] S. Malik *et al.*, "Paper-based sensors: affordable, versatile, and emerging analyte detection platforms," *Anal. Methods*, vol. 16, no. 18, pp. 2777–2809, 2024, doi: 10.1039/D3AY02258G.

[11] Y. Liu, S. Lu, Z. Zhang, Z. Yang, X. Cui, and G. Liu, "Printable biosensors towards next-generation point-of-care testing: paper substrate as an example," *Lab. Chip*, vol. 23, no. 15, pp. 3328–3352, 2023, doi: 10.1039/D3LC00038A.

[12] P. Srikrishnarka, J. Haapasalo, J. P. Hinestroza, Z. Sun, and Nonappa, "Wearable Sensors for Physiological Condition and Activity Monitoring," *Small Sci.*, vol. 4, no. 7, p. 2300358, Jul. 2024, doi: 10.1002/smsc.202300358.

[13] Patro, Pramoda, et al. "A hybrid approach estimates the real-time health state of a bearing by accelerated degradation tests, Machine learning." *2021 Second International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)*. IEEE, 2021.

[14] N. Colozza, V. Mazzaracchio, and F. Arduini, "Paper-Based Electrochemical (Bio)Sensors for the Detection of Target Analytes in Liquid, Aerosol, and Solid Samples," *Annu. Rev. Anal. Chem.*, vol. 17, no. 1, pp. 127–147, Jul. 2024, doi: 10.1146/annurev-anchem-061522-034228.

[15] S. Benjamin, F. De Lima, V. Nascimento, G. De Andrade, and R. Oriá, "Advancement in Paper-Based Electrochemical Biosensing and Emerging Diagnostic Methods," *Biosensors*, vol. 13, no. 7, p. 689, Jun. 2023, doi: 10.3390/bios13070689.

[16] Palaniradja, K., and V. Balasubramanian. "Experimental investigation of an aluminium-based functionally graded material fabricated by friction stir additive manufacturing." *Materials Research Express* 13.1 (2026): 016504.

[17] X. Zhang, J. Li, Z. Cai, L. Zhao, and C. Liu, "Deep Learning-Based Signal Quality Assessment for Wearable ECGs," *IEEE Instrum. Meas. Mag.*, vol. 25, no. 5, pp. 41–52, Aug. 2022, doi: 10.1109/MIM.2022.9832823.

[18] R. He, H. Liu, Y. Niu, H. Zhang, G. M. Genin, and F. Xu, "Flexible Miniaturized Sensor Technologies for Long-Term Physiological Monitoring," *Npj Flex. Electron.*, vol. 6, no. 1, p. 20, Mar. 2022, doi: 10.1038/s41528-022-00146-y.

[19] G. Backiyalakshmi, U. Snekhalatha, and A. L. Salvador, "Recent advancements in non-invasive wearable electrochemical biosensors for biomarker analysis – A review," *Anal. Biochem.*, vol. 692, p. 115578, Sep. 2024, doi: 10.1016/j.ab.2024.115578.

[20] S. R. Stahlschmidt, B. Ulfenborg, and J. Synnergren, "Multimodal deep learning for biomedical data fusion: a review," *Brief. Bioinform.*, vol. 23, no. 2, p. bbab569, Mar. 2022, doi: 10.1093/bib/bbab569.

[21] S. Chen *et al.*, "Wearable flexible microfluidic sensing technologies," *Nat. Rev. Bioeng.*, vol. 1, no. 12, pp. 950–971, Sep. 2023, doi: 10.1038/s44222-023-00094-w.

[22] S. R. S. Pour *et al.*, "Microfluidic-Based Non-Invasive Wearable Biosensors for Real-Time Monitoring of Sweat Biomarkers," *Biosensors*, vol. 14, no. 1, p. 29, Jan. 2024, doi: 10.3390/bios14010029.