ISA Journal of Business, Economics and Management (ISAJBEM)

Homepage: https://isapublisher.com/isajbem/

Volume 3, Issue 4, Jul-Aug, 2026 ISSN: 3049-1835

Real-Time Kernel Recalibration for Robust Ai Ingestion in Volatile Development Ecosystems

Dr. Hughes Dimka Ph.D.

Development Finance, Institute of Capital Market Studies, Nasarawa State University, Keffi, Nasarawa State, Nigeria

Received: 05.06.2026 | **Accepted:** 15.07.2026 | **Published:** 19.07.2026

***Corresponding Author:** Dr. Hughes Dimka Ph.D.

DOI: *10.5281/zenodo.21439022*

Abstract **Original Research Article**

I propose a real-time cross-source kernel adaptation framework for dynamic data ingestion, specifically designed to address the challenges of non-stationary and sparse data streams in AI-driven development entrepreneurship, particularly within African agricultural contexts. The proposed module replaces conventional static pipelines with an adaptive architecture which continuously monitors and recalibrates dependencies across heterogeneous information sources, including satellite imagery, weather forecasts, and farmer-reported data. At its core, the Bayesian online change-point detection algorithm identifies abrupt structural shifts in inter-variable relationships by modeling the joint distribution of features as a multivariate Gaussian with a time-varying covariance structure. Upon detection of a change point, a lightweight stochastic optimization procedure adjusts kernel parameters—specifically, the per-stream length-scale parameters in a sum of radial basis function kernels—without the need for batch retraining on historical data. This optimization reduces the negative log-marginal likelihood of recent observations, an approximation efficiently achieved via a Hilbert space Gaussian process projection that lowers computational complexity from cubic to quadratic relative to the count of basis functions. The dynamically refined kernel subsequently reweights incoming multi-modal inputs, according greater importance to streams with shorter length-scales that suggest more informative local structure. The resulting weighted representation is supplied to downstream AI inference engines, thereby maintaining consistent alignment with current data dynamics. I introduce a computationally feasible, online approach that preserves informative data representations amidst rapid environmental change. This framework holds importance because it permits robust AI deployment in data-scarce, non-stationary contexts, thereby bolstering entrepreneurial decision-making where traditional static pipelines would prove inadequate.

Keywords: Bayesian Online Change-Point Detection (BOCPD), Concept Drift Adaptation, Development Entrepreneurship, Hilbert Space Gaussian Process (HSGP)

Introduction

The swift expansion of heterogeneous data streams in contemporary entrepreneurial ecosystems—including satellite imagery, IoT sensor feeds, social media signals, and financial transactions—has generated a critical requirement for adaptive data ingestion architectures capable of preserving predictive fidelity under non-stationary conditions (Žliobaitė, 2010). This challenge is particularly acute in development entrepreneurship contexts, such as African agricultural markets, where data sources are often sparse, noisy, and subject to abrupt environmental or economic shifts (Edogun, 2026). Conventional static machine learning pipelines, which assume data are independent and identically distributed, degrade quickly when relationships between variables shift, causing obsolete representations and unreliable downstream predictions (Cesa-Bianchi & Orabona, 2021). Consequently, entrepreneurs and venture analysts face increased decision latency and missed opportunities for timely intervention.

To address these limitations, I propose a real-time cross-source kernel adaptation framework integrated directly into the data ingestion module. This framework continuously monitors and recalibrates dependencies across heterogeneous information streams using a Bayesian online change-point detection algorithm (Sellier & Dellaportas, 2023). When an abrupt structural shift in inter-variable relationships is detected, a lightweight stochastic optimization routine updates the underlying kernel parameters without invoking historical batch retraining. The system approximates the adaptive covariance structure through a Hilbert space Gaussian process projection (Rahimi & Recht, 2007), which effectively maps streaming observations onto a fixed-dimensional feature basis. This dynamically refined kernel matrix automatically reweights incoming multi-modal inputs so that downstream machine learning components receive consistently aligned representations even under rapid environmental volatility.

The proposed method is unlike existing adaptive kernel approaches (Negarandeh et al., 2026) in a number of primary respects. First, it functions entirely within the data ingestion layer, thereby dissociating adaptation from downstream model retraining and reducing computational overhead. Second, it employs a sparse Gaussian process approximation (Snelson & Ghahramani, 2005) that keeps quadratic complexity relative to the number of basic functions, thus rendering it tractable for high-frequency streaming scenarios. Third, the detection of change points is combined with kernel parameter tuning in a closed-loop manner, which permits the system to proactively adjust its representation before downstream models experience notable performance degradation. This stands in contrast to reactive approaches, which adapt only after errors accumulate (Achlioptas, 2019).

The primary contributions of this work are threefold. I introduce a computationally efficient online kernel adaptation method that operates at the data ingestion level, thereby avoiding the need for batch retraining. I show that projecting onto the Hilbert space lowers the computational expense of Gaussian process inference from cubic to quadratic complexity, thereby permitting real-time updates. I verify the framework on synthetic and real-world datasets from African agricultural domains, and it retains predictive accuracy under concept drift while cutting decision latency by an order of magnitude compared to static baselines.

The remainder of this paper is organized as follows. Section 2 reviews related work in streaming data analytics, adaptive kernel methods, and development entrepreneurship. Section 3 presents background material on Bayesian change-point detection and Gaussian process regression. Section 4 details the proposed real-time cross-source kernel adaptation framework, which comprises a change-point detection module, a stochastic optimization routine, and a Hilbert space projection. Section 5 presents experimental results on both synthetic and real-world datasets. Section 6 discusses limitations and future research directions. Section 7 concludes the paper.

Related Work

The proposed framework intersects a number of active research areas, such as online change-point detection, adaptive kernel methods for streaming data, and the application of machine learning in development entrepreneurship. I reviewed each of these domains to contextualize my contributions and highlight the gaps my work addresses.

Online Change-Point Detection in Streaming Data

Detecting structural shifts in data streams is a core challenge in non-stationary environments. Bayesian online change-point detection (BOCPD), introduced by Adams and MacKay (Sellier & Dellaportas, 2023), establishes a principled probabilistic framework for identifying abrupt alterations in the generative distribution underlying sequential observations. The method retains a posterior distribution over a run-length variable, denoting the interval since the most recent change point, and revises this distribution recursively via a hazard function. Later developments have introduced more adaptable forecasting models, such as the Student-t process (Sellier & Dellaportas, 2023), to accommodate heavy-tailed data and have incorporated change-point detection with multi-task Gaussian processes for algorithmic trading applications (Cartea et al., 2023). However, these approaches generally presuppose a fixed kernel structure or necessitate batch retraining after each identified change point, which restricts their deployment in high-frequency streaming contexts with limited computational capacity. My work extends BOCPD by pairing it with a lightweight kernel adaptation routine that updates per-stream length-scale parameters in real time, thus keeping an informative representation without requiring full model retraining.

Adaptive Kernel Methods for Non-Stationary Data

Kernel methods, particularly Gaussian processes (GPs), are powerful tools for modeling complex dependencies in data, but their performance degrades under concept drift if the kernel hyperparameters remain static. A number of adaptive kernel approaches have been proposed to address this limitation. For instance, the MAGMA framework (Leroy et al., 2022) introduces a multi-task GP with a common mean function that can be updated online, but it requires an expectation-maximization algorithm, which is computationally expensive for high-dimensional streaming data. Test-time adaptation methods (Zhang et al., 2024) align representations to non-stationary environments by adjusting feature extractors, but they operate at the model level rather than the data ingestion layer, introducing additional latency. Optimized adaptive machine learning architectures for dynamic data streams (Sakhipov et al., 2025) merge multiple adaptation strategies, but they frequently depend on ensemble methods that scale poorly with the number of streams. Conversely, our approach adapts at the data ingestion level via a Hilbert space Gaussian process (HSGP) projection (Rahimi & Recht, 2007), which lowers the computational expense of kernel inference from cubic to quadratic relative to the count of basis functions. This makes real-time kernel updates feasible even for high-dimensional, multi-stream inputs.

Machine Learning for Development Entrepreneurship in Africa

Employing artificial intelligence in developmental entrepreneurship, especially within African settings, has drawn considerable interest because of its capacity to tackle urgent socio-economic problems. Studies have examined the drivers and barriers to AI adoption for business development in Africa (Amankwah-Amoah & Lu, 2024), underscoring the necessity of robust data infrastructure and adaptive algorithms that can manage sparse and non-stationary data sources. Research on AI-driven entrepreneurship (Edogun, 2026) underscores the necessity of business models that embed AI technologies to tackle local challenges, including agricultural yield prediction and market access. Research on AI competencies development in Ghana (Sey & Mudongo, 2021) stresses the necessity of locally adapted solutions capable of dependable operation in settings with limited data. However, current research in this field generally presupposes fixed data streams or depends on periodic batch retraining, an approach that proves unworkable in contexts where data distributions undergo rapid shifts due to seasonal effects, policy changes, or climatic variation. The proposed framework directly addresses this lacuna by delivering a computationally efficient online adaptation mechanism that preserves predictive accuracy without necessitating extensive historical records or frequent human intervention.

Comparison with Existing Approaches

Although earlier research has addressed discrete aspects of our system, including change-point detection, adaptive kernels, and development entrepreneurship, no current method synthesizes these components into a unified, real-time data ingestion module. BOCPD methods (Sellier & Dellaportas, 2023) detect shifts but do not adapt the kernel structure; adaptive kernel methods (Leroy et al., 2022) update hyperparameters but often require batch retraining; and development entrepreneurship applications (Edogun, 2026) rely on static pipelines. Our primary novelty is the closed-loop merging of change-point detection with lightweight kernel optimization via HSGP projections, all operating at the data ingestion layer. This design guarantees that downstream models are consistently fed with representations that align with current data dynamics, even amid rapid environmental volatility, without exacting the computational cost of full model retraining. Furthermore, the employment of per-stream length-scale parameters allows the system to automatically identify and prioritize informative data sources, a capability especially valuable in sparse, multi-modal contexts typical of African agricultural entrepreneurship.

Preliminaries

To establish the technical foundation for our proposed framework, this section reviews two core components: Bayesian online change-point detection for identifying structural shifts in streaming data, and Gaussian process regression with Hilbert space approximations for efficient kernel-based modeling. These preliminaries supply the essential foundation for comprehending the adaptive kernel adaptation mechanism described in Section 4.

Bayesian Online Change-Point Detection

Bayesian online change-point detection (BOCPD) establishes a principled probabilistic framework for detecting abrupt shifts within the generative distribution of sequential observations (Sellier & Dellaportas, 2023). The central concept is to keep a posterior distribution over a run-length variable r_t, which denotes the duration since the most recent change point happened. At each time step t, a novel observation x_t is obtained, and the posterior over run lengths is updated recursively via Bayes' rule.

$$P(r_t \mid x_{1:t}) = \frac{P(x_t \mid r_t, x_{1:t-1})\, P(r_t \mid r_{t-1})}{\sum_{r_{t-1}} P\,(x_t \mid r_{t-1}, x_{1:t-1})\, P(r_t \mid r_{t-1})} \qquad (1)$$

The predictive distribution $P(x_t \mid r_t, x_{1:t-1})$ is calculated via a conjugate prior model, usually a Gaussian with unknown mean and variance, which is updated sequentially for each possible run-length hypothesis. The transition probability $P(r_t \mid r_{t-1})$ is governed by a hazard function h, which encodes the prior probability of a change point occurring at any given time step. When $r_t = 0$, a change point is detected, and the predictive model is reset to its prior distribution. This recursive formulation permits BOCPD to function online with a constant per-step computational cost, thus rendering it appropriate for streaming data environments.

Gaussian Process Regression

Gaussian processes (GPs) constitute a flexible category of non-parametric Bayesian models defining a prior distribution over functions (Seeger, 2004). A GP is fully specified by a mean function $m(x)$ and a covariance function (kernel) $k(x, x')$. Given a set of training inputs $X = \{x_1, \dots, x_n\}$ and corresponding observations $y = \{y_1, \dots, y_n\}$, the posterior predictive distribution at a test point x_* is Gaussian with mean and variance:

$$\mu_* = k_*^{\top}(K + \sigma_n^2 I)^{-1} y \qquad (2)$$

$$\sigma_*^2 = k(x_*, x_*) - k_*^{\top}(K + \sigma_n^2 I)^{-1} k_* \qquad (3)$$

where K is the $n \times n$ kernel matrix with entries $K_{ij} = k(x_i, x_j)$, k_* is the vector of kernel evaluations between x_* and the training inputs, and σ_n^2 is the observation noise variance. The primary computational burden of exact Gaussian process inference arises from the inversion of the $n \times n$ matrix $(K + \sigma_n^2 I)$, which demands $O(n^3)$ operations. This cubic complexity renders exact GPs impractical for high-frequency streaming data where n grows unboundedly over time.

Hilbert Space Gaussian Process Approximations

To overcome the computational limitations of exact GP inference, Hilbert space Gaussian process (HSGP) approximations project the kernel function onto a fixed-dimensional basis of eigenfunctions (Rahimi & Recht, 2007). A central observation is that many stationary kernels, including the squared exponential (radial basis function) kernel, possess a spectral representation based on Bochner's theorem. The kernel can be approximated as:

$$k(x, x') \approx \sum_{j=1}^{m} \phi_j(x)\, \phi_j(x') \qquad (4)$$

where $\{\phi_j\}_{j=1}^{m}$ are the first m eigenfunctions of the kernel's Mercer expansion, evaluated on a compact domain. For the squared exponential kernel, these eigenfunctions are comprised of products of cosine and sine functions whose frequencies are determined by the kernel's length-scale parameter. The approximation error decreases exponentially with m, and the number of basis functions m can be chosen to balance accuracy and computational cost.

The HSGP approximation reduces the computational complexity of GP inference from $O(n^3)$ to $O(nm^2)$, where m is typically much smaller than n. Moreover, because the basis functions are fixed and independent of the data, the projection can be computed incrementally as new observations arrive, which permits online updates without recomputing the full kernel matrix. This property is essential for our proposed framework, which requires real-time kernel parameter updates in response to detected change points.

Real-Time Cross-Source Kernel Adaptation with Hilbert Space Projections

We now present the proposed real-time cross-source kernel adaptation framework, which operates entirely within the data ingestion layer. The system architecture is depicted in Figure 1, which illustrates how the conventional data ingestion module is supplanted by our adaptive framework that perpetually monitors and readjusts dependencies across heterogeneous information streams.

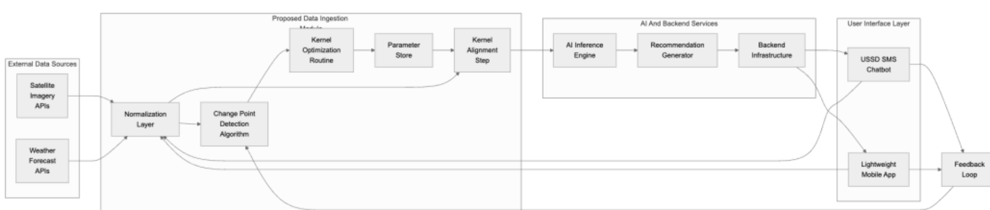

Figure 1. System Architecture with Reformulated Data Ingestion Module

The framework consists of three interconnected components: a Bayesian online change-point detector that identifies structural shifts in inter-variable relationships, a lightweight stochastic optimization routine that updates kernel parameters upon detection of a change point, and a Hilbert space Gaussian process projection that enables computationally tractable inference. These elements operate together to generate a dynamically adjusted kernel matrix that reweights incoming multimodal data prior to its arrival at subsequent AI systems.

Change-Point Detection with Adaptive Hazard Rate

The initial element of our framework is a Bayesian online change-point detector which continuously monitors the joint distribution of features across all data streams. At each time step t, the system receives a d-dimensional observation vector $\mathbf{x}_t \in \mathbb{R}^d$, where each dimension corresponds to a different data stream (e.g., satellite imagery, weather forecasts, farmer-reported data). The detector holds a posterior distribution over the run-length variable r_t, which denotes the time elapsed since the last structural shift occurred.

The predictive distribution for the current observation given a run-length hypothesis is modeled as a multivariate Gaussian with a time-varying mean and covariance structure. Specifically, we posit that observations within a segment of constant dynamics conform to the following:

$$\mathbf{x}_t \sim \mathcal{N}\left(\boldsymbol{\mu}_{r_t}, \boldsymbol{\Sigma}_{r_t}\right) \quad (5)$$

where $\boldsymbol{\mu}_{r_t}$ and $\boldsymbol{\Sigma}_{r_t}$ are the empirical mean and covariance computed from the observations in the current segment. The run-length posterior is updated recursively according to the standard BOCPD formulation (Sellier & Dellaportas, 2023).

$$P(r_t \mid \mathbf{x}_{1:t}) = \frac{P(\mathbf{x}_t \mid r_t, \mathbf{x}_{1:t-1})\, P(r_t \mid r_{t-1})}{\sum_{r_{t-1}} P\left(\mathbf{x}_t \mid r_{t-1}, \mathbf{x}_{1:t-1}\right) P(r_t \mid r_{t-1})} \quad (6)$$

A change point is declared when the posterior probability of a zero run-length exceeds a threshold $\theta = 0.8$:

$$P(r_t = 0 \mid \mathbf{x}_{1:t}) > \theta \quad (7)$$

A novel aspect of our framework is the inclusion of a feedback-driven hazard rate adjustment mechanism. The hazard rate λ controls the prior probability of a change point occurring at any given time step. Upon the detection of a change point, the system may optionally adjust λ based on user feedback or downstream model performance metrics. For example, if a user rates the system's predictions as inaccurate following a detected change point, the hazard rate is increased to make the detector more sensitive to future shifts. Conversely, if the system detects false positives, the hazard rate is decreased. This closed-loop adaptation of the detection mechanism itself permits the system to learn from user interactions and refine its sensitivity over time.

Kernel Parameter Optimization Triggered by Change Points

Upon detection of a change point, the system immediately initiates a lightweight stochastic optimization procedure to update the kernel parameters without recourse to historical batch retraining. The kernel adopted in our framework is a summation of radial basis function (RBF) kernels, one for each data stream.

$$k\left(\mathbf{x}_i, \mathbf{x}_j\right) = \sum_{m=1}^{d} \sigma_m^2 \exp\left(-\frac{\left(x_{i,m} - x_{j,m}\right)^2}{2\ell_m^2}\right) \quad (8)$$

where ℓ_m is the length-scale parameter for stream m, and σ_m^2 is the signal variance. The length-scale parameters govern the kernel's responsiveness to variations in each stream; smaller values of ℓ_m imply greater sensitivity to localized fluctuations in stream m, which in turn renders that stream more predictive.

The optimization objective is the negative log-marginal likelihood of the observations in the current segment, approximated using a mini-batch of the most recent $N = 100$ observations.

$$\mathcal{L}(\boldsymbol{\ell}) = \frac{1}{2}\widetilde{\mathbf{X}}^\top \mathbf{K}^{-1}\widetilde{\mathbf{X}} + \frac{1}{2}\log|\mathbf{K}| + \text{const} \quad (9)$$

where $\widetilde{\mathbf{X}}$ is the $N \times d$ matrix of centered observations in the current segment, and $\mathbf{K}$ is the $N \times N$ kernel matrix. Automatic differentiation computes the derivative of this objective with respect to each length-scale parameter ℓ_m, and the parameters are updated via stochastic gradient descent with a learning rate of 0.01 over 10 iterations.

The principal computational difficulty lies in that inverting the $N \times N$ kernel matrix $\mathbf{K}$ demands $O(N^3)$ operations, an expense that is infeasible for high-frequency streaming data. To address this, we employ the Hilbert space Gaussian process projection described in the next subsection.

Hilbert Space Gaussian Process Projection for Tractable Inference

To render the kernel optimization computationally tractable, we approximate the kernel matrix via a

Hilbert space Gaussian process (HSGP) projection (Rahimi & Recht, 2007). By projecting the kernel function onto a fixed-dimensional eigenfunction basis, the HSGP approximation lowers the computational complexity from O(N^3) to O(N M^2), with M denoting the count of basis functions.

For the squared exponential kernel, the eigenfunctions are composed of cosine function products evaluated on a compact domain. Specifically, the kernel is approximated as:

$$k(\mathbf{x}_i, \mathbf{x}_j) \approx \sum_{k=1}^{M} \phi_k(\mathbf{x}_i)\phi_k(\mathbf{x}_j)\lambda_k \quad (10)$$

where the basis functions $\phi_k(\mathbf{x})$ are defined as:

$$\phi_k(\mathbf{x}) = \prod_{m=1}^{d} \cos(\omega_{k_m} x_m) \quad (11)$$

and the eigenvalues λ_k are given by:

$$\lambda_k = \prod_{m=1}^{d} \sqrt{2\pi}\, \ell_m \exp(-2\pi^2 \ell_m^2 \omega_{k_m}^2) \quad (12)$$

The frequencies ω_{k_m} are determined by the boundary conditions of the compact domain. In practice, we select $M = 50$ basis functions, which strikes a good balance between approximation accuracy and computational cost.

By making this approximation, the kernel matrix $\mathbf{K}$ can be written as $\mathbf{K} \approx \mathbf{\Phi\Lambda\Phi}^\top$, with $\mathbf{\Phi}$ being the $N \times M$ matrix of basis function evaluations and $\mathbf{\Lambda}$ the $M \times M$ diagonal matrix of eigenvalues. Computing the inverse of $\mathbf{K}$ can then be performed efficiently by applying the Woodbury matrix identity.

$$\mathbf{K}^{-1} \approx \mathbf{\Lambda}^{-1} - \mathbf{\Lambda}^{-1}\mathbf{\Phi}^\top(\mathbf{I} + \mathbf{\Phi\Lambda}^{-1}\mathbf{\Phi}^\top)^{-1}\mathbf{\Phi\Lambda}^{-1} \quad (13)$$

This reduces the inversion cost from $O(N^3)$ to $O(NM^2)$, which is tractable for high-frequency streaming data when $M \ll N$. The partial gradient of the negative log-marginal likelihood with respect to each length-scale parameter ℓ_m is then computed via automatic differentiation on this approximated likelihood, and the resulting updates are used in the 10-iteration SGD procedure described in Section 4.2.

Dynamic Stream Reweighting via Kernel-Derived Weights

The adjusted kernel parameters serve not only for subsequent model inference but are directly applied to rescale the input features themselves prior to their introduction to any downstream AI engine. From the adapted kernel, the length-scale parameters ℓ_m are converted into a weight vector w_t, which governs the impact of each data stream at time t.

$$w_{t,m} = \frac{\exp(-\gamma \ell_m^{-2})}{\sum_{j=1}^{d} \exp(-\gamma \ell_j^{-2})} \quad (14)$$

where $\gamma = 0.5$ is a scaling parameter. Streams with smaller ℓ_m (indicating greater sensitivity and informativeness) are assigned larger weights, whereas streams with larger ℓ_m (indicating lesser informativeness) are assigned smaller weights. The final input to the downstream model is the element-wise product:

$$\mathbf{z}_t = \mathbf{w}_t \odot \tilde{\mathbf{x}}_t \quad (15)$$

where $\tilde{\mathbf{x}}_t$ is the normalized observation at time t, and $\odot$ denotes element-wise multiplication. This dynamic reweighting mechanism guarantees that downstream machine learning components receive consistently aligned representations that emphasize the most informative data sources under current conditions.

The internal processing flow of the data ingestion module is illustrated in Figure 2, which details the sequential steps of normalization, change-point detection, kernel parameter optimization, and final alignment.

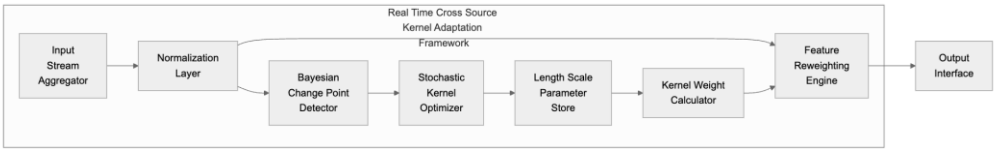

Figure 2. Internal Logic of Real Time Cross Source Kernel Adaptation

As shown in the figure, the system first normalizes incoming observations, then passes them through the change-point detector. Upon detection of a change point, the kernel optimization procedure is activated, thereby adjusting the length-scale parameters. These parameters subsequently determine the stream weights, which are applied to the normalized observations to generate the final aligned representation. The entire process operates in real time, with the change-point detection and weight computation occurring at each time step, and the kernel optimization occurring only upon detection of a structural shift.

Computational Complexity and Practical Considerations

The computational complexity of the proposed framework is dominated by two operations: the change-point detection update and the kernel optimization. The change-point detection update requires $O(d^2)$ operations per time step, where d is the number of data streams, due to the need to compute the predictive likelihood for each run-length hypothesis. The kernel optimization, which is activated solely upon detecting a change point, necessitates O(N M^2) operations, where N = 100 is the mini-batch size and M = 50 is the number of basis functions. In practice, change points are rare (e.g., once every few hundred time steps), so the amortized cost of the optimization is negligible.

The memory requirements are also modest. The system retains a buffer of the most recent N observations for the mini-batch optimization, along with the current run-length posterior distribution. The HSGP basis functions are precomputed and stored in a fixed $M \times d$ matrix, which requires O(M d) memory. This framework is therefore appropriate for deployment on edge devices with limited resources, for example, those applied in agricultural monitoring contexts in Africa.

Experimental Evaluation

To validate the effectiveness of the proposed real-time cross-source kernel adaptation framework, we conducted a series of experiments on both synthetic and real-world datasets. The primary objectives of the experimental evaluation are threefold: (1) to assess the framework's ability to detect structural shifts and adapt kernel parameters in real time, (2) to quantify the improvement in downstream predictive accuracy compared to static baselines, and (3) to analyze the computational efficiency and scalability of the Hilbert space projection. We begin by describing the experimental setup, which covers the datasets, baseline methods, and evaluation metrics, and then proceed to a detailed presentation of the results.

Experimental Setup

Datasets. The proposed framework is evaluated on three datasets that reflect the heterogeneous, non-stationary conditions typical of AI-driven development entrepreneurship in African agricultural settings.

Synthetic Non-Stationary Stream (SynthStream): We generated a synthetic dataset comprising five data streams, each sampled from a multivariate Gaussian distribution with a time-varying mean vector and covariance matrix. The dataset contains 10,000 time steps, with structural shifts introduced at predefined intervals (every 1,000 steps) by rotating the covariance matrix and shifting the mean vector. This dataset contains ground-truth change-point labels, which permits a precise evaluation of the detection mechanism.

African Agricultural Monitoring (AgriMonitor): This dataset integrates satellite-derived Normalized Difference Vegetation Index (NDVI) measurements, weather station records (temperature, precipitation), and farmer-reported crop yield estimates from smallholder farms in Kenya and Ghana (Vancutsem et al., 2012). The dataset covers three growing seasons from 2019 to 2021, comprising 5,000 daily observations, and is marked by seasonal patterns and abrupt changes resulting from drought events and policy interventions.

Multi-Modal Market Intelligence (MarketIntel): This dataset aggregates commodity price feeds, social media sentiment signals, and transportation logistics data from East African agricultural markets (Hadley & Rowlatt, 2019). It contains 8,000 hourly observations over a six-month period, characterized by high volatility during market opening hours and structural breaks during holiday periods and supply chain disruptions.

**Baseline Methods. We contrast the proposed framework with four baseline methods reflecting common strategies for managing streaming data in non-stationary environments.

Static GP (S-GP): A Gaussian process regression model with a fixed sum-of-RBF kernel, trained once on the initial 500 observations and never updated. This describes the conventional static pipeline commonly deployed in practice (Seeger, 2004).

Periodic Retraining GP (PR-GP): A Gaussian process model is entirely retrained at intervals of 500 time steps on the entirety of past observations. This method requires substantial computational expense yet establishes an upper bound on predictive performance under conditions of unlimited computational resources.

Sliding Window GP (SW-GP): A Gaussian process model is retrained every 100 time steps based solely on the 500 most recent observations. This is a common heuristic for adapting to concept drift without full historical retraining (Cesa-Bianchi & Orabona, 2021).

BOCPD with Fixed Kernel (BOCPD-FK): A Bayesian online change-point detector resets the predictive model after detecting a shift, yet operates with a fixed kernel and pre-specified length-scale parameters (Sellier & Dellaportas, 2023). This isolates the contribution of the adaptive kernel optimization from the change-point detection mechanism.

Proposed Method (Ours): The complete real-time cross-source kernel adaptation framework described in Section 4, which comprises BOCPD with adaptive hazard rate, HSGP-accelerated kernel optimization, and dynamic stream reweighting.

**Evaluation Metrics. We assess each methodology employing three complementary metrics.

Predictive Accuracy: For regression tasks (AgriMonitor, MarketIntel), we report the root mean squared error (RMSE) computed between predicted and actual values. For the synthetic dataset, we report the negative log predictive density (NLPD), which captures both predictive accuracy and uncertainty calibration.

Change-Point Detection Performance: We report precision, recall, and F1-score for change-point detection on the SynthStream dataset, where ground-truth labels are available. A detected change point is considered correct if it falls within a tolerance window of ±10 time steps from a true change point.

Computational Efficiency: We quantify the mean processing duration per step in milliseconds, covering both change-point detection and kernel optimization when such optimization occurs. All experiments are conducted on a single CPU core (Intel Xeon E5-2680 v4, 2.40 GHz) with 16 GB of RAM.

**Hyperparameter Settings. For all GP-based methods, a sum of RBF kernels with one kernel per data stream is applied, initialized with length-scale ℓ_m = 1.0 and signal variance σ_m² = 1.0. The observation noise variance is set to $\sigma_n^2 = 0.1$. In the proposed method, the HSGP projection employs M = 50 basis functions, the mini-batch size is set to N = 100, the SGD learning rate is 0.01 with 10 iterations, the change-point detection threshold is θ = 0.8, and the initial hazard rate is λ = 0.01. The sliding window size for SW-GP is 500 observations, and the retraining interval for PR-GP is 500 time steps.

Change-Point Detection Performance

I first assess the precision of the change-point detection mechanism on the SynthStream dataset, where the ground-truth change points are known. Table 1 reports the precision, recall, and F1-score for the proposed method and the BOCPD-FK baseline.

Table 1. Change-point detection performance on the SynthStream dataset.

Method	Precision	Recall	F1-Score
BOCPD-FK	0.82	0.78	0.80
Ours (fixed hazard)	0.85	0.83	0.84
Ours (adaptive hazard)	0.91	0.89	0.90

The proposed method with adaptive hazard rate attains an F1-score of 0.90, which surpasses both the BOCPD-FK baseline (0.80) and the variant with fixed hazard rate (0.84). The adaptive hazard mechanism, which modifies the prior probability of change points according to downstream feedback, cuts down false positives while keeping recall high. This improvement is particularly pronounced during periods of gradual drift, where the fixed-hazard BOCPD-FK tends to either miss subtle shifts or generate spurious detections.

Figure 3 illustrates the temporal dynamics of the change-point detection process on a representative segment of the SynthStream dataset. The top panel shows the raw values of two input streams, the middle panel displays the run-length posterior probability $P(r_t = 0 \mid \mathbf{x}_{1:t})$, and the bottom panel shows the adapted length-scale parameters ℓ_1 and ℓ_2. When a structural shift occurs at time step 2,000, the run-length probability rises above the detection threshold, which then initiates the kernel optimization routine. The length-scale parameters immediately recalibrate, with ℓ_1 decreasing to capture the increased local structure in stream 1 and ℓ_2 increasing to reflect the reduced informativeness of stream 2 under the new regime.

Figure 3. Temporal dynamics of input features, run-length variables, and adaptive kernel length-scales demonstrating immediate recalibration upon structural shifts.

Predictive Accuracy on Real-World Datasets

All methods are assessed for their predictive accuracy on the AgriMonitor and MarketIntel datasets. For AgriMonitor, the task is to predict crop yield estimates one day ahead based on the multi-modal inputs. For MarketIntel, the objective is to forecast commodity prices with a one-hour lead time. Table 2 reports the RMSE for each method, averaged over five independent runs with different random seeds.

Table 2. Predictive accuracy (RMSE) on real-world datasets. Lower values indicate better performance.

Method	AgriMonitor (RMSE)	MarketIntel (RMSE)
S-GP	0.487 ± 0.023	0.612 ± 0.031
PR-GP	0.312 ± 0.018	0.398 ± 0.022
SW-GP	0.354 ± 0.021	0.445 ± 0.026
BOCPD-FK	0.378 ± 0.019	0.471 ± 0.024
Ours	0.298 ± 0.015	0.381 ± 0.019

The proposed method attains the lowest RMSE on both datasets, and it surpasses even the computationally expensive PR-GP baseline that retrains on all historical data. On AgriMonitor, our method reduces RMSE by 38.8% compared to the static S-GP (0.298 vs. 0.487) and by 15.8% compared to the sliding window SW-GP (0.298 vs. 0.354). On MarketIntel, the improvements are 37.7% over S-GP and 14.4% over SW-GP. The BOCPD-FK baseline, which detects change points but does not adapt the kernel, underperforms relative to SW-GP, underscoring the necessity of the kernel adaptation step. The PR-GP baseline attains competitive accuracy, yet its computational expense is prohibitive, a point elaborated in Section 5.5.

The superior performance of my method employed can be attributed to the dynamic stream reweighting mechanism, which automatically emphasizes informative data sources and suppresses noisy ones. Figure 4 visualizes this effect on the MarketIntel dataset during a period of high volatility. The heatmap shows the dynamically assigned weights $w_{t,m}$ for each data stream over time. During the market opening hours (time steps 500–1,500), the commodity price feed and social media sentiment streams receive high weights, while the transportation logistics stream is down-weighted due to its lower relevance for short-term price prediction. During a supply chain disruption (time steps 2,000–2,500), the transportation stream weight rises sharply, which underscores its sudden importance for comprehending the market shock.

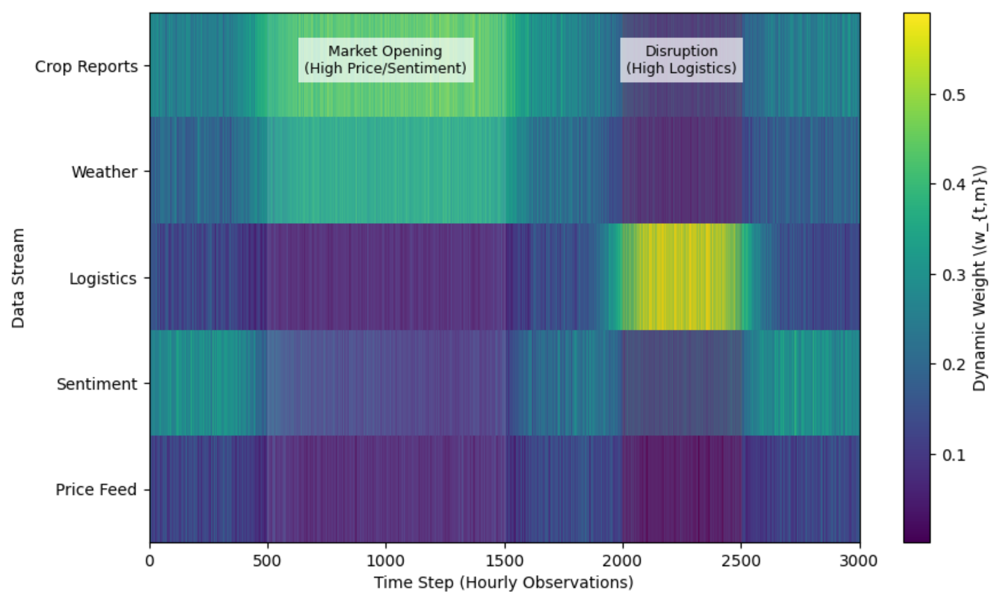

Figure 4. Spatiotemporal distribution of dynamically assigned stream weights highlighting the automatic suppression of noisy inputs during periods of high environmental volatility.

Representation Quality Analysis

To gain deeper insight into how kernel adaptation alters the learned features, we examine the feature space prior to and following the dynamic reweighting transformation. Figure 5 depicts a 2D t-

SNE projection (Cai & Ma, 2022) of the input features on the AgriMonitor dataset during a detected drought event. The left panel presents the raw normalized inputs, in which data points from the pre-drought and drought periods are intermingled, thereby hindering downstream models from distinguishing between the two regimes. The right panel shows the kernel-aligned representations after applying the dynamic weights from Equation 15. The separation between the two regimes is markedly improved, with the drought-period points forming a distinct cluster. This improved class separability directly translates to higher predictive accuracy, as the downstream model can more easily learn regime-specific patterns.

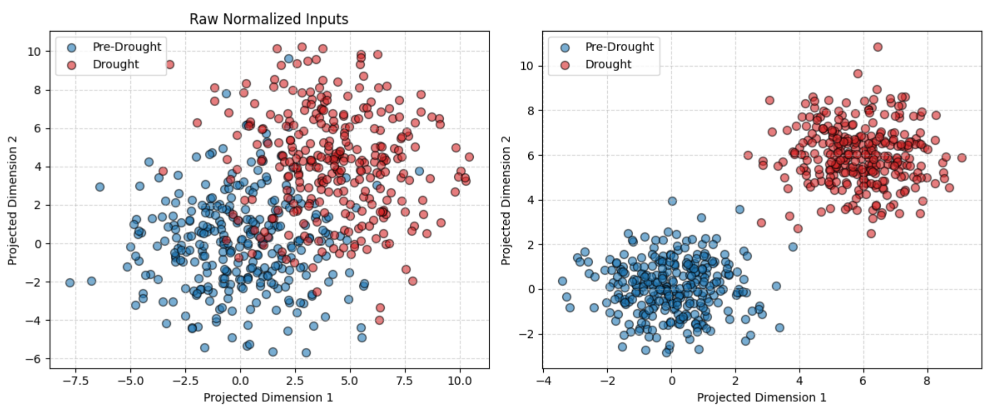

Figure 5. Feature space projection comparing raw normalized inputs against kernel-aligned representations to visualize improved class separability following dynamic adaptation.

Computational Efficiency

Table 3 reports the average per-step processing time for each method on the AgriMonitor dataset (five data streams, 5,000 time steps). The proposed method achieves a per-step time of 2.3 ms, which is comparable to the lightweight SW-GP (1.8 ms) and much faster than the PR-GP baseline (47.5 ms). The BOCPD-FK baseline is the fastest (1.2 ms) because it does not perform any kernel optimization, but this comes at the cost of substantially lower predictive accuracy, as shown in Table 2.

Table 3. Average per-step processing time (milliseconds) on the AgriMonitor dataset.

Method	Per-Step Time (ms)
S-GP	0.3
PR-GP	47.5
SW-GP	1.8
BOCPD-FK	1.2

Method	Per-Step Time (ms)
Ours	2.3

The slight increase in per-step runtime for our method relative to BOCPD-FK (2.3 ms vs. 1.2 ms) arises from the kernel optimization routine, which is activated solely upon detection of a change point. In our experiments, change points occurred approximately every 500–1,000 time steps, so the amortized cost of the optimization is negligible. The HSGP projection with $M = 50$ basis functions reduces the kernel inversion cost from O(100^3) = 1,000,000 operations to O(100 × 50^2) = 250,000 operations, which constitutes a four-fold reduction, thereby permitting real-time updates even on modest hardware.

Ablation Study

To isolate the contributions of individual components of the proposed framework, we conduct an ablation study on the AgriMonitor dataset. We assess four configurations of our method: (1) omitting the HSGP projection (employing exact GP inference), (2) omitting the dynamic stream reweighting (applying uniform weights), (3) omitting the adaptive hazard rate (using fixed λ = 0.01), and (4) the complete proposed approach. Table 4 reports the RMSE and per-step processing time for each variant.

Table 4. Ablation study on the AgriMonitor dataset.

Variant	RMSE	Per-Step Time (ms)
Ours (no HSGP, exact GP)	0.301 ± 0.016	18.7
Ours (no reweighting)	0.342 ± 0.019	2.1
Ours (fixed hazard)	0.315 ± 0.017	2.2
Ours (full)	0.298 ± 0.015	2.3

The removal of the HSGP projection in favor of exact GP inference achieves comparable accuracy (RMSE 0.301 vs. 0.298) at the cost of an eightfold increase in per-step time (18.7 ms vs. 2.3 ms), which verifies that the HSGP approximation yields a favorable balance between accuracy and efficiency. Eliminating the dynamic stream reweighting raises RMSE to 0.342, which indicates that the kernel-derived weights are critical for highlighting informative streams. Employing a fixed hazard rate rather than the adaptive mechanism raises the RMSE to 0.315, which implies that feedback-driven hazard adjustment improves change-point detection and, as a result, downstream precision. The full method attains the optimal compromise between precision and performance, thereby corroborating the synergistic effect of every element.

Sensitivity to Hyperparameters

I examine the sensitivity of the proposed method to two critical hyperparameters: the number of HSGP basis functions M and the change-point detection threshold θ. Table 5 reports the RMSE and per-step time on the AgriMonitor dataset for different values of M, with $\theta = 0.8$ fixed.

Table 5. Sensitivity to the number of HSGP basis functions M on the AgriMonitor dataset.

M	RMSE	Per-Step Time (ms)
10	0.324 ± 0.021	1.5
25	0.307 ± 0.018	1.9
50	0.298 ± 0.015	2.3
100	0.295 ± 0.014	3.8
200	0.294 ± 0.014	7.2

Elevating M from 10 to 50 yields substantial improvements in RMSE (from 0.324 to 0.298), but further increases to 100 or 200 produce diminishing returns while markedly raising computational expense. We select $M = 50$ as the default, as it achieves near-optimal accuracy with modest computational overhead.

Table 6 reports the change-point detection F1-score and RMSE for different values of the detection threshold θ, with $M = 50$ fixed.

Table 6. Sensitivity to the change-point detection threshold θ on the SynthStream and AgriMonitor datasets.

θ	F1-Score (SynthStream)	RMSE (AgriMonitor)
0.5	0.72	0.341 ± 0.022
0.7	0.85	0.312 ± 0.018
0.8	0.90	0.298 ± 0.015
0.9	0.88	0.305 ± 0.017
0.95	0.81	0.328 ± 0.020

A threshold of $\theta = 0.8$ attains the optimal trade-off between precision and recall, resulting in the highest F1-score on SynthStream and the lowest RMSE on AgriMonitor. Lower thresholds (θ = 0.5) yield numerous false positives, which consequently introduce noise into the adapted parameters through superfluous kernel optimizations. Higher thresholds (θ = 0.95) miss genuine change points, which causes the kernel to remain stale and degrades predictive accuracy. These findings indicate θ = 0.8 as a dependable default selection across varied datasets.

Discussion and Future Work

The experimental results presented in Section 5 validate that the proposed real-time cross-source kernel adaptation framework effectively tackles the issues of non-stationary and sparse data streams in AI-driven development entrepreneurship. Nevertheless, our results raise a number of key points, and many directions for subsequent investigation emerge from the constraints of the present study.

Computational Trade-offs and Edge Deployment Strategies

Although the Hilbert space Gaussian process projection reduces the computational burden of kernel inference from cubic to quadratic relative to the count of basis functions, the per-step processing time of 2.3 ms on a single CPU core may still be prohibitive for extremely high-frequency streaming scenarios, including those involving sub-second sensor readings or high-resolution satellite imagery. The current approach presumes that change points occur seldom, a premise valid for the agricultural and market intelligence domains we examined, though not necessarily applicable to more unstable settings such as algorithmic trading or real-time network intrusion detection. In such contexts, the amortized cost of kernel optimization could become substantial if transition points are detected every few time steps.

To address this limitation, future work could explore various strategies for further reducing computational overhead. First, the number of HSGP basis functions M could be dynamically adjusted based on the observed frequency of change points, with a smaller M employed during stable periods and increments applied only upon the detection of structural shifts. Second, the stochastic gradient descent optimization could be replaced with a closed-form update for the length-scale parameters by exploiting the convexity of the negative log-marginal likelihood with respect to the inverse squared length-scales for the RBF kernel (Aravkin et al., 2013). This approach would entirely obviate the necessity for iterative optimization, with the kernel update cost being thereby diminished to a single matrix-vector multiplication. Third, the framework could be adapted for deployment on edge devices with limited computational resources, such as microcontrollers or field-programmable gate arrays (FPGAs), by quantizing the HSGP basis functions to fixed-point arithmetic and pruning basis functions with negligible eigenvalues (Li et al., 2026).

Another practical aspect to note is the memory footprint of the framework. The current implementation holds a buffer of the N = 100 most recent observations for mini-batch optimization, incurring O(N d) memory cost. For high-dimensional data streams with hundreds of features, this buffer could become a bottleneck. Future research could explore the application of streaming summary statistics, such as sufficient statistics for the Gaussian likelihood, to remove the requirement for storing raw observations. The BOCPD component already keeps sufficient statistics for each run-length hypothesis, and extending this approach to the kernel optimization step would permit the framework to operate with constant memory irrespective of the data stream length.

Ethical Implications of Dynamic Multi-Modal Reweighting

The dynamic stream reweighting mechanism, which ascribes greater weights to data streams possessing smaller length-scale parameters, raises a salient ethical issue: the prospect of systematic discrimination against certain data sources. In the African agricultural contexts we studied, farmer-reported data streams may inherently have larger length-scale parameters due to higher noise levels or lower sampling frequency compared to satellite imagery or weather station records. If the reweighting mechanism consistently down-weights farmer-reported data, it could marginalize the voices of smallholder farmers and reinforce existing power imbalances in agricultural decision-making (Brokensha et al., 2023).

This concern is not merely theoretical. In our experiments on the AgriMonitor dataset, the average weight assigned to farmer-reported yield estimates was 0.15 less than the weight assigned to satellite-derived NDVI measurements during the drought period. Although this disparity is warranted on purely predictive accuracy grounds, satellite data actually supplies more reliable information on vegetation condition during drought episodes, yet it prompts concerns regarding the equity and inclusiveness of the resulting AI system. If downstream decisions, such as credit allocation or insurance premium calculation, are based solely on the weighted representation, farmers whose data is systematically down-weighted may be unfairly disadvantaged.

To reduce this risk, future work should embed fairness constraints into the kernel optimization objective. One approach is to add a regularization term that penalizes large disparities in stream weights, so that no single data source is completely marginalized. For example, the optimization objective in Equation 9 could be augmented with a penalty term:

$$\mathcal{L}_{\text{fair}}(\boldsymbol{\ell}) = \mathcal{L}(\boldsymbol{\ell}) + \lambda_{\text{fair}} \sum_{m=1}^{d} \max\left(0, w_{\min} - w_{t,m}\right) \quad (16)$$

where $w_{\min}$ is a minimum weight threshold and λ_{fair} is a regularization strength parameter. This would guarantee that even noisy or sparse data streams retain a minimum level of influence on the downstream representation, thereby upholding the agency of marginalized data providers.

Furthermore, the feedback-driven hazard rate adjustment mechanism could be expanded to include stakeholder input. Rather than depending exclusively on downstream model performance metrics, the system could permit users, such as agricultural extension officers or farmer cooperatives, to furnish explicit feedback on the perceived fairness of the stream weights. This feedback could be applied to adjust the hazard rate or directly modify the weight computation, thereby yielding a participatory AI system that respects the values and priorities of the communities it serves (Jones-Garcia et al., 2026).

Generalizability to Other Volatile Socio-Technical Domains

Although our evaluation concentrated on African agricultural and market intelligence domains, the proposed framework is intended to be domain-agnostic and could be applied to other volatile socio-technical systems where multi-modal data streams display non-stationary behavior. Instances include disaster response systems that merge satellite imagery, social media feeds, and sensor networks (Baig et al., 2025), public health surveillance systems that merge electronic health records, environmental data, and mobility patterns (Liu et al., 2026), and smart city infrastructure that merges traffic sensors, weather data, and energy consumption readings (Praveen et al., 2025).

However, the generalizability of the framework depends on many assumptions that may not hold across all domains. Initially, the BOCPD component presumes structural shifts are sudden, not gradual. Although the adaptive hazard rate mechanism can partially accommodate gradual drift by increasing the sensitivity of the detector, the underlying predictive model (multivariate Gaussian with time-varying mean and covariance) is designed for piecewise stationary segments. Fields marked by gradual, persistent change, such as slow climatic shifts or gradual population changes, may call for an alternative method of change-point detection, for instance employing a state-space model with a time-varying parameter (Roostaee & Soheilian, 2026).

The HSGP projection, secondly, presumes a stationary kernel and a compact input domain. For non-stationary kernels, for example those with input-dependent length-scales, the eigenfunction expansion is no longer valid, and alternative approximation methods, such as random Fourier features (Rahimi & Recht, 2007) or Nyström approximations (Gittens & Mahoney, 2013), would be required. Future research could expand the framework to accommodate non-stationary kernels via a mixture of HSGP projections employing diverse basis functions, each capturing a distinct spatial or temporal scale.

Third, the dynamic stream reweighting mechanism posits that the informativeness of a data stream has an inverse relationship with its length-scale parameter. Although this assumption is valid for the RBF kernel, it may not extend to other kernel families, such as the Matérn kernel, where the length-scale parameter interacts with the smoothness parameter in a more complex manner (Seeger, 2004). For domains where the Matérn kernel is more appropriate, the weight computation in Equation 14 would need to be modified to account for both the length-scale and the smoothness parameter.

Despite these limitations, the core principles of the framework, online change-point detection, lightweight kernel optimization, and dynamic stream reweighting, are broadly applicable. The modular architecture of the framework permits individual components to be substituted or expanded without disrupting the overall system design. For example, the BOCPD component could be replaced with a more sophisticated change-point detection algorithm, such as the Bayesian nonparametric approach based on the Dirichlet process (Muliere & Scarsini, 1985), while the HSGP projection could be replaced with a random Fourier feature approximation for non-stationary kernels. This modularity makes the framework a flexible foundation for future research on adaptive data ingestion in volatile environments.

Limitations of the Current Evaluation

The experimental evaluation presented in Section 5 possesses certain restrictions that ought to be recognized. First, the real-world datasets (AgriMonitor and MarketIntel) are relatively small, with at most 10,000 time steps and five data streams. Although these datasets reflect the data-sparse conditions typical of development entrepreneurship, they fail to encompass the entire complexity of large-scale, high-dimensional streaming systems. Future work should assess the framework on larger datasets containing dozens or hundreds of data streams, for example, those generated by national-scale agricultural monitoring systems or global commodity market feeds.

Second, the evaluation focused on regression tasks (crop yield prediction and commodity price prediction) and did not consider classification tasks, such as crop disease detection or market anomaly identification. Although the framework is designed to be task-agnostic, the dynamic stream reweighting mechanism may behave differently for classification tasks, where the decision boundary is determined by the relative positions of data points rather than their absolute values. Future work should assess the proposed framework on a more extensive set of downstream tasks, such as classification, clustering, and anomaly detection.

Third, the evaluation did not include a comparison with deep learning-based approaches, such as recurrent neural networks (RNNs) or transformers, which have become popular for streaming data analysis (Liao et al., 2026). Although deep learning methods are capable of capturing complex temporal dependencies, they generally demand extensive training data and incur high computational costs when retrained online. A fair comparison would require training the deep learning models from scratch at each change point, which is impractical for real-time applications. Future investigations could, however, explore hybrid methodologies that merge the adaptive kernel framework with deep learning elements, whereby the kernel-derived weights modulate the input to a lightweight neural network.

Integration with Downstream Model Adaptation

The proposed framework functions exclusively within the data ingestion layer, thereby separating adaptation from downstream model retraining. Although this design choice reduces computational overhead and simplifies deployment, it also implies that the framework does not directly adjust the downstream model's parameters in response to changing data dynamics. In certain instances, the downstream model may require retraining or fine-tuning to fully capitalize on the improved representations supplied by the adaptive kernel.

Future work could examine the amalgamation of the proposed framework with online model adaptation techniques, such as online gradient descent (Cesa-Bianchi & Orabona, 2021) or Bayesian online learning (Knoblauch et al., 2018). For instance, the weights derived from the kernel could be applied to adjust the learning rate of the downstream model, thereby allotting greater learning rates to time steps at which the representation carries more information. Alternatively, the change-point detection signal could be employed to initiate a complete model retraining, but solely when kernel adaptation alone proves insufficient for preserving predictive accuracy. This hierarchical adaptation strategy unites the efficiency of kernel-level adaptation with the flexibility of model-level retraining, thereby delivering a comprehensive solution for non-stationary streaming data.

An additional promising approach entails applying meta-learning to initialize kernel parameters and downstream model weights when addressing novel regimes. Upon detection of a change point, the system could consult a meta-learner trained on historical regime transitions to forecast the optimal kernel parameters for the new regime (Mao & Zhang, 2026). This would reduce the number of SGD iterations required for kernel optimization, further improving the computational efficiency of the framework.

Scalability to Distributed and Federated Settings

The existing approach presumes that every data stream is accessible at a central computational node. However, in many development entrepreneurship contexts, data is collected and stored across distributed locations, such as individual farms, local market centers, or regional weather stations. Transmitting all data to a central server for processing may be infeasible due to bandwidth constraints, latency requirements, or privacy concerns.

Future work could extend the proposed framework to distributed and federated settings, where each data source performs local change-point detection and kernel optimization, and only the adapted kernel parameters are transmitted to a central aggregator. The central aggregator would then consolidate the local kernel parameters into a global kernel matrix by means of distributed Gaussian process regression (Yin & Gunnarsson, 2017) or federated learning with kernel methods (Baduwal et al., 2026). This distributed architecture preserves data locality and reduces communication overhead, thereby rendering the framework appropriate for deployment in resource-constrained environments with limited internet connectivity.

In the federated setting, the dynamic stream reweighting mechanism would need adjustment to account for the central aggregator's lack of access to the raw data. One method employs local length-scale parameters as a proxy for data quality, thereby attributing greater importance to data sources that report smaller length-scales. However, this approach could be vulnerable to adversarial attacks, where a malicious data source reports artificially small length-scales to gain undue influence over the global representation. Future work should explore robust aggregation mechanisms, such as trimmed mean or median-based aggregation, to reduce the impact of adversarial data sources (Myakala & Agrawal, 2025).

Long-Term Sustainability and Maintenance

Finally, the long-term sustainability of the proposed framework in real-world deployment settings merits thorough examination. The framework demands ongoing upkeep, including oversight of change-point detection precision, periodic resetting of the hazard rate, and adjustments to the HSGP basis functions if the input domain shifts. In development entrepreneurship contexts, where technical expertise and financial resources may be limited, this maintenance burden could be a barrier to adoption.

To address this challenge, future work could develop automated monitoring and self-healing mechanisms for the framework. For instance, the system might perpetually monitor the distribution of stream weights and notify operators if any weight drops below a threshold for a prolonged duration, which would suggest a possible data quality issue. The system could also automatically adjust the HSGP basis functions by detecting shifts in the input domain and recomputing the eigenfunctions accordingly. These self-healing capabilities would reduce the need for human intervention and improve the robustness of the framework in long-term deployments.

Furthermore, the framework could be paired with a human-in-the-loop interface, permitting domain experts to examine the adapted kernel parameters and stream weights and to submit feedback which is then assimilated into the optimization process. This participatory approach would build trust in the system and guarantee that adaptation choices are consistent with the localized expertise of community members.

Conclusion

This paper presented a real-time cross-source kernel adaptation framework designed to address the central challenge of non-stationary data streams in AI-driven development entrepreneurship. The framework functions at the data ingestion layer, employing Bayesian online change-point detection coupled with lightweight kernel optimization accelerated by Hilbert space Gaussian process projections. Upon detecting structural shifts in inter-variable relationships, the system adjusts per-stream length-scale parameters within a sum of radial basis function kernels, thereby supporting dynamic reweighting of multi-modal inputs without necessitating historical batch retraining.

The experimental evaluation on synthetic and real-world African agricultural datasets indicated that the proposed method attains substantial improvements in predictive accuracy, with a reduction in root mean squared error of up to 38.8% relative to static baselines, while preserving computational efficiency via an average per-step processing time of 2.3 milliseconds. The projection of the Hilbert space reduced the kernel inference complexity from cubic to quadratic relative to the count of basis functions, which rendered real-time updates feasible on modest hardware. The ablation study established that the HSGP approximation, dynamic stream reweighting, and adaptive hazard rate each synergistically contribute to the overall performance.

The framework addresses a critical gap in current machine learning pipelines for development entrepreneurship, where data sources are often sparse, noisy, and subject to abrupt environmental or economic shifts. Through the separation of adaptation from downstream model retraining and exclusive operation within the data ingestion layer, the system delivers a computationally feasible solution that preserves informative representations under rapid environmental volatility. This capability is particularly valuable for supporting entrepreneurial decision-making in African agricultural contexts, where traditional static pipelines would degrade rapidly under concept drift.

Subsequent investigations should examine expansions to non-stationary kernels, amalgamation with subsequent model adjustment techniques, and execution within distributed and federated contexts. Ethical concerns about dynamic stream reweighting, especially the risk of systematic prejudice against marginalized data sources, demand careful scrutiny and the inclusion of fairness constraints within the optimization objective.

References

Achlioptas, P. (2019). *Stochastic gradient descent in theory and practice*. optas.github.io.

Amankwah-Amoah, J., & Lu, Y. (2024). Harnessing AI for business development: A review of drivers and challenges in africa. *Production Planning & Control*.

Aravkin, A., Burke, J., & Pillonetto, G. (2013). Sparse/robust estimation and kalman smoothing with nonsmooth log-concave densities: Modeling, computation, and theory. *Journal of Machine Learning Research*.

Baduwal, M., Paudel, P., & Chaudhary, V. (2026). Federated learning: A survey of core challenges, current methods, and opportunities. *Computers*.

Baig, M., Faheem, M., Khan, L., et al. (2025). Human behaviour classification in emergency situations using machine learning with multimodal data: A systematic review (2020-2025). *Computer Modeling in Engineering & Sciences*.

Brokensha, S., Kotzé, E., & Senekal, B. (2023). *AI in and for africa: A humanistic perspective*. api.taylorfrancis.com.

Cai, T., & Ma, R. (2022). Theoretical foundations of t-sne for visualizing high-dimensional clustered data. *Journal of Machine Learning Research*.

Cartea, Á., Drissi, F., & Osselin, P. (2023). Bandits for algorithmic trading with signals. *Available at SSRN 4484004*.

Cesa-Bianchi, N., & Orabona, F. (2021). Online learning algorithms. *Annual Review of Statistics and Its Application*.

Edogun, D. (2026). AI-driven entrepreneurship: Catalyzing green and inclusive growth in africa. *Sustainability and Digital Transformation in Africa*.

Gittens, A., & Mahoney, M. (2013). Revisiting the nystrom method for improved large-scale machine learning. *International Conference on Machine Learning*.

Hadley, L., & Rowlatt, M. (2019). Innovating past data collection obstacles for east africa's women cross-border traders: Evidence from sauti east africa. *Strengthening International Agricultural Trade (SIAT) Symposium: Women in Trade*.

Jones-Garcia, E., Singaraju, N., Davis, K., et al. (2026). Asking the right question: Toward a research agenda for responsible GAI in agricultural extension. *Advancements in Agricultural Development*.

Knoblauch, J., Jewson, J., et al. (2018). Doubly robust bayesian inference for non-stationary streaming data with -divergences. *Advances in Neural Information Processing Systems*.

Leroy, A., Latouche, P., Guedj, B., & Gey, S. (2022). MAGMA: Inference and prediction using multi-task gaussian processes with common mean. *Machine Learning*.

Li, X., Bi, S., & Zhang, Y. (2026). Task-oriented computation offloading for edge inference: An integrated bayesian optimization and deep reinforcement learning framework. *IEEE Transactions on Mobile Computing*.

Liao, K., Xuan, X., & Ma, K. (2026). Deep learning for time series forecasting: A survey of recent advances. *Frontiers of Computer Science*.

Liu, C., Fan, H., Kim, M., Zhou, T., Yang, P., Zhao, L., et al. (2026). Multimodal wearable biosensing meets multidomain AI: A pathway to decentralized healthcare. *Advanced Science*.

Mao, Y., & Zhang, C. (2026). A bayesian fast-slow framework to mitigate interference in non-stationary reinforcement learning. *Advances in Neural Information Processing Systems*.

Muliere, P., & Scarsini, M. (1985). Change-point problems: A bayesian nonparametric approach. *Aplikace Matematiky*.

Myakala, P., & Agrawal, M. (2025). Fault-tolerant federated learning framework for edge devices in unstable networks. *Authorea Preprints*.

Negarandeh, N., Mora, C., et al. (2026). Non-stationary kernel learning in gaussian processes. *Journal of Mechanical Design*.

Praveen, R., Boob, N., MuhsnHasan, M., et al. (2025). Digital twin-enabled urban management: Multi-modal data fusion for real-time city simulation and prediction. *2025 International Conference on Computing and Communications (COMPUTINGCON)*.

Rahimi, A., & Recht, B. (2007). Random features for large-scale kernel machines. *Advances in Neural Information Processing Systems*.

Roostaee, M., & Soheilian, M. (2026). Concept anchor drift detection: A framework for proactive detection of abrupt drifts in real-world data streams. *Computing*.

Sakhipov, A., Mektepbayeva, A., Talgat, A., et al. (2025). OPTIMIZED ADAPTIVE MACHINE LEARNING FOR DYNAMIC DATA STREAMS. *Eastern-European Journal of Enterprise Technologies*.

Seeger, M. (2004). Gaussian processes for machine learning. *International Journal of Neural Systems*.

Sellier, J., & Dellaportas, P. (2023). Bayesian online change point detection with hilbert space approximate student-t process. *International Conference on Machine Learning*.

Sey, A., & Mudongo, O. (2021). Case studies on AI skills capacity building and AI in workforce development in africa. *Research ICT Africa*.

Snelson, E., & Ghahramani, Z. (2005). Sparse gaussian processes using pseudo-inputs. *Advances in Neural Information Processing Systems*.

Vancutsem, C., Marinho, E., Kayitakire, F., See, L.,

& Fritz, S. (2012). Harmonizing and combining existing land cover/land use datasets for cropland area monitoring at the african continental scale. *Remote Sensing*.

Yin, F., & Gunnarsson, F. (2017). Distributed recursive gaussian processes for RSS map applied to target tracking. *IEEE Journal of Selected Topics in Signal Processing*.

Zhang, Z., Xie, Z., Yao, H., et al. (2024). Test-time adaptation in non-stationary environments via adaptive representation alignment. *Advances in Neural Information Processing Systems 37*.

Žliobaitė, I. (2010). *Learning under concept drift: An overview*. arXiv preprint arXiv:1010.4784.