# Drift-Aware Machine Learning for Precision Agriculture: A Survey on Concept Drift Detection and Adaptive Learning in IoT Environmental Monitoring

*Amit Kumar Bindal[1], Dr. Shashi Kant Gupta [2]*

[1,2] Computer Science and Engineering, Lincoln University College, Petaling Jaya, Selangor, Malaysia

Email ID: pdf.bindalamit@lincoln.edu.my

**Abstract**

The use of algorithms and internet of things sensor networks in precision agriculture continues to grow, However as conditions in the field change over time, so does the original model. This paper will focus on the evaluation of techniques for drift detection and adaptation in precision agriculture systems. In particular, we will examine hybrid approaches that combine various types of sensors in agriculture. We will analyze different techniques for detecting changes in models, including adaptive windowing methods, statistics based approaches (several N-dimensional extensions of Kolmogorov-Smirnov tests), and incremental updating methods (such as using online sequential extreme learning machines). One of the primary contributions of this paper is an integrated drift detection system that is referred to as the Hybrid Adaptive Window (ADWIN) system. This system uses N-Dimensional KS Testing along with Hoeffding's bounds in order to allow for real-time updates of models. As part of this survey, we will also share our analysis about issues of scaling for computation, long term benchmarking against new models and federated deployment challenges. Finally, we will identify several significant limitations associated with edge computing efficiency and privacy and recommend several significant opportunities for future work related to the creation of intelligent, distributed agrometeorological monitoring systems capable of providing timely measurements while conserving limited resources.

**Keywords:** Concept drift, Precision agriculture, IoT sensors, Edge Computing, Adaptive windowing.

## Introduction

The advent of IoT sensor networks represents a significant transformation in how we monitor our environment due to their ability to generate high-quality data continuously and store it in a usable format, thus allowing the use of data-driven approaches to manage resources better over time [1]. The use of IoT sensor networks is prevalent within the precision farming or 'smart' farming realm, where a variety of multi-modal types of environmental data are being collected, including soil moisture levels, temperature, humidity, pH, light intensity, and rainfall, among other variables, so that effective crop management decisions can be made based on these data sources [2]. Furthermore, these data sources are being used in conjunction with machine learning algorithms to predict crop yields, detect anomalies, and create automated irrigation schedules [3]. By 2024, the global market for smart agriculture IoT will be worth approximately $13.5 billion, while by 2034, it is expected to surpass $40 billion (4), indicating that many technological advances are occurring at an accelerated pace [4].

However, One major challenge impeding the effectiveness of machine learning (ML) based agricultural systems is that they suffer from "concept drift." As data streams evolve over time, their statistical characteristics change as well, which results in a divergence between the joint distributions of input and target variables from what was originally used to train the model — thereby degrading the model's predictive ability [5]. In agriculture, both seasonal changes and climate changes can contribute to drift; also, changing fertilization and/or irrigation regimes may also create a drift in an agricultural context [6]. Therefore, any static models developed based on past data will eventually become out of date because of these changes in input/output relationships. So, to ensure reliable ML-based agricultural systems, we will need adaptive learning systems capable of automatically detecting recent changes in distributions of variable pairings (also known as concept drift) and adaptively learning how to compensate for them [7].

ADWIN introduced the adaptive windowing (ADWIN) algorithm which adapts the size of its sliding window based on changes in the input and gives statistically backed assurances for the false positive and false negative rates of the system [8]. Although ADWIN is widely used today as a primary means of detecting drifts, it was originally designed to detect drift for a single variable. Because many agricultural sensors concurrently have their data correlated (e.g., affecting multiple variables), this creates challenges when trying to detect drift in such settings. These challenges have spurred research involving multiple variables with the goal of testing the hypothesis of and using multidimensional versions of the ADWIN algorithm; many researchers have started experimenting with the distribution test known as the Kolmogorov-Smirnov statistic to measure drift across multiple sensors; however, no research has yet focused on these types of approaches with the same size/scale as is commonly found in resource-constrained IoT devices [9].

**Literature Survey**

**2.1 Concept Drift Detection in Data Streams**

In the last decade, the area of detecting concept drift in streaming data has made a significant advancement; i.e., there are now many more algorithms for detecting distributional changes than there were at the start of this decade. Christoph Raab et al created the Reactive Robust Soft Learning Vector Quantization (RRSLVQ) algorithm, which combines a Kolmogorov-Smirnov (K-S) Test-based drift detector and prototype adaptation mechanisms to solve the stability vs plasticity issue by maintaining stable performance during drift and allowing fast adaptations to new concepts. The K-S test provides a non-parametric approach to comparing empirical cumulative distributions of two data windows (one from recent data, the other from historical data), thus it provides an empirical basis for drift detection without needing to make any assumptions about the parametric distributional forms of either window of data [10].

Viktor et al introduced the installation of IoT-based agricultural systems, where edge devices need to deal with high-frequency, low-latency, and high-volume sensor data streams from environmental sensors, is critical to the operational performance of modern agriculture systems. The IKS algorithm can be used to detect unsupervised drift in real-time without the need for true label-based verification in field deployments where the costs associated with label validation are prohibitively expensive and/or time-consuming. Therefore, IKS provides a computationally efficient solution for performing a K-S-based statistical test and achieving a respective K-S test statistic result [11].

**2.2 Drift Detection in Image and Visual Data Stream**

Qunag-Tien Tran et al. discussed while multi-variate temporal sensor data is still considered to be the primary way in which agricultural monitoring is conducted, advances made around computer vision-based drift detection are providing much needed methodological knowledge for transfer to other types of environmental monitoring. The author developed a method to monitor for feature-space drift in image data streams using intermediate neural network representations, rather than raw pixel distributions of images. The author's approach uses the penultimate layer activations of a convolutional network to create a reference distribution, which incoming batches can be compared against using a statistical divergence measure. Monitoring latent representations of input data rather than the actual raw input data by embedding those input measurements into a lower dimensional feature space prior to drift analysis is therefore a direct method of applying the findings of this paper to the agricultural systems being studied [12].

The context-aware drifting detection framework, developed by Oliver Cobb et al., expands traditional two-sample homogeneity testing methods into conditional distribution-based settings. Since operations typically exhibit contextually dependent correlations that contradict i.i.d. assumptions, their context-based calculation of maximum conditional mean discrepancy can identify drift across populations without being affected by trends of prevalence. This essential capability will be beneficial to agronomists for analytics of various times of crop growing phases or soil zones; there are variations in baseline distributions across the context zones, and so the drift will need to be detected for each context without confusing the variation in context with actual changes in the distribution. The framework's applicability to problems with ImageNet-level complexity provides a basis for confirming its potential scalability for the high-dimensional sensor arrays used in many of today's sophisticated farms [13].

Yoshihiro Okawa, and Kenichi Kobayashi developed the boundary shrinking method of concept drift detection (CDDBS) is an approach to detecting changes in real-world distributions based on the distance between sets of historical and recent examples. For example, if there were no random removable points placed into the training set for the inspector, then inspecting will tend to become "more aggressive" and detect drift earlier compared to inspectors that have had random removable points placed into the training set. An alternative to using one inspector to detect classification score changes due to drifts would be to use an ensemble of inspectors. In other words, you could use multiple inspectors as a statistical means of assessing the extent of drift. Using an ensemble of inspectors, you could develop inspector-specific or crop-specific drift detector models. Therefore, this approach would support heterogeneous agricultural environments for harvest-related activities (e.g. sugar beets vs. potatoes) and heterogeneous sensor-based environments (e.g. ultrasound vs. radar). Consequently, there would be no need for the actual systems to be isolated as they would all be connected in some fashion by the agnostic nature of the model used to characterize the compound change that occurred over time [14].

### 2.3 Adaptive Learning and Model Update Strategies

Fuccellaro has demonstrated an extensive doctoral study on drift detection, update and correction and has produced the Gaussian Split Detector (GSD) for unsupervised drift detection from image streams.

The GSD models use Gaussian Mixture Models to define distributions, and they monitor shifts in optimal decision boundaries of an ensemble of simple splits that are specifically targeting real concept drift as opposed to virtual drift which will not have a significant effect on the classification boundaries. The reliance of the method on Gaussian assumptions and its focus on shifts in decision boundaries as opposed to complete changes in whole distributions provides a cost-effective solution for edge deployments [15].

### 2.4 Federated and Edge- Intelligent Agricultural Systems

Research has been sparked using adaptive learning in agricultural environments with limited resources. In their systematic review of federated learning, Mahdi et al. pointed out several obstacles related to concept drift--among them, differences between end user devices, asynchronous update situations, as well as the distinction between local and global drift types. The authors show that regular drift detection methods do not fit the criteria for federated learning, as data collected by the devices has no independent or identically distributed (i.i.d.) characteristics from other clients, and data cannot be shared because privacy laws prohibit sharing raw data with others. A key area where federated drift detection can improve, specifically in agriculture IoT networks where many farms experience diverse types of soil and climate conditions, is finding ways to offer federated drift detection methods that align both local and global distributions of drift without compromising privacy [16].

## Conclusion

This paper provides a review of techniques and methods that can be used to identify and adapt to changes in the real-time drift of data during the monitoring of environmental conditions in agriculture. Therefore, it focuses specifically on hybrid approaches to combining distribution-testing with multiple sensors (testing equipped with multiple sensors) and the use of lightweight online learning. We demonstrate how static machine learning models in precision agriculture are subject to concept drift, which is caused by seasonal changes in the environment, climate change, and human-induced disturbances. As a result, the need for adaptive frameworks arises in order to detect the distributional changes that occur due to these external forces and update predictions without requiring a full retraining of the model(s) being used. Our analysis reveals the following three convergent pathways: (I) the evolution of univariate drift detection methods to multivariate drift detection methods via ADWIN extensions, variants of the Kolmogorov-Smirnov method, and methods based on the maximum mean discrepancy for the joint monitoring of correlated/associated parameters; (II) the increasing separation of virtual drift from real drift through use of context-aware detectors (e.g., lightweight classifiers) and ensemble-based techniques (e.g., ensemble classifiers) that reduce the boundaries of both; and (III) the integration of drift detection with online sequential learning techniques (e.g., online self-tuning anomaly detector) to facilitate the efficient deployment of these methods on the edge of computing architectures.

The Hybrid ADWIN framework combines N-dimensional K-S testing with Hoeffding's Bound and online self-tuning anomaly detection (OL-OS-ELM) and provides significant benefits in practice, with an improvement in accuracy from 85.86% to 97.29%. However, challenges still remain in achieving computational scaling for microcontroller class devices, the lack of long-term multi-season benchmarking, the lack of standardized streaming protocols for agricultural data, and the need for more federated architectures that maintain the privacy and confidentiality of agricultural data across multiple farms.

## Future Scope

The following areas should receive priority attention in future research. First, there is a need to design low weight drift detectors that maintain statistical rigor while considering memory and energy limitations of battery-powered agricultural sensors. Second, we need to create open benchmark datasets of environmental streams that capture multiple seasons, multiple geographic regions, and multiple crop types to allow for rigorous comparative evaluations of these systems. Finally, by integrating explainable artificial intelligence (XAI) with drift detection frameworks, farmers will be able to have greater confidence in their sensors and algorithms due to receiving interpretable alerts of reasons for sensor-specific triggers and model updates. Furthermore, combining edge intelligence with federated learning and adaptive drift management will provide privacy-preserving, scalable ecosystems for maintaining predictive accuracy under changing environmental conditions. In addition, long-term seasonal stability studies extending several years instead of a few weeks; developing standardized evaluation methodologies focused on agricultural streaming scenarios; and investigating availability of neuromorphic computing architectures for ultra-low power edge inference represent important pathways for advancing resilient, deployable Agricultural monitoring systems.

## References


1. Gurram Bhavani, Arpita Baronial, Sarangam Kodati and M. Dhasaratham, “IoT-Based Environmental Monitoring System with Machine Learning for Accurate and Real-time Data Analysis”, Library Progress International, Vol. 32, No. 3, Jul-Dec 2024, pp 20548-20560.
2. Abhinav Sharm, Arpit Jain, Prateek Gupta and Vinay Chowdary, “Machine Learning Applications for Precision Agriculture: A Comprehensive Review”, IEEE Access, Vol. 9, January 2021, pp. 4843-4873.
3. Ravesa Akhter and Shabir Ahmad Sofi, “Precision agriculture using IoT data analytics and machine learning”, Journal of King Saud University- Computer and Information Sciences, Vol. 34, No. 8, September 2022, pp. 5602-5618.
4. IoT Precision Agriculture Sensors Market https://market.us/report/iot-precision-agriculture-sensors-market/, Report ID: 138264, January 2025.
5. João Gama, Indrė Žliobaitė, Albert Bifet, Mykola Pechenizkiy, Abdelhamid Bouchachia, “A survey on concept drift adaptation”, Vol. 46, No. 4, ACM Computing Survey, March 2014, pp. 1-37.
6. Hezal Lopes and Prashant Nitnaware, “Real Time Drift Detection and Adaptation Using Hybrid ADWIN in Agricultural Environmental Monitoring System, International Journal of Electrical and Electronics and Telecommunications, Vol. 14 No. 5, July 2025, pp. 313-322.
7. Fabian Hinder, Valerie Vaquet and Barbara Hammer, “One or two things we know about concept drift-a survey on monitoring in evolving environments. Part-A: detecting concept drift”, Frontiers in Artificial Intelligence, Vol. 7, June 2024, DOI: https://doi.org/10.3389/frai.2024.1330257.
8. Albert Bifet and Ricard Gavaldà, “Learning from Time-Changing Data with Adaptive Windowing”, Proceedings of the 2007 SIAM International Conference on Data Mining, 2007, pp. 443-448.

9. Daniel Lukats, Oliver Zielinski, Axel Hahn and Frederic Stah, " A benchmark and survey of fully unsupervised concept drift detectors on real-world data streams", International Journal of Data Science and Analytics, Vol. 19, 2025, pp. 1-31.
10. Christoph Raab, Moritz Heusinger and Frank-Michael Schleif, " Reactive Soft Prototype Computing for concept Drift Streams", Vol. 416, November 2020, pp. 340-351.
11. Viktor Losing, Barbara Hammer and Heiko Wersing, "KNN Classifier with Self Adjusting Memory for Heterogeneous Concept Drift", IEEE 16th International Conference on Data Mining (ICDM), December 2016, pp. 291-300, **DOI:** 10.1109/ICDM.2016.0040
12. Quang-Tien Tran, Nhien-An-Le-Khac and Michela Bertolotto,"Concept drift detection in image data stream: a survey on current literature, limitations and future directions", Artificial Intelligence Review, Vol. 59, No. 33, December 2025.
13. Oliver Cobb and Arnaud Van Looveren, "Context-Aware Drift Detection", Proceedings of the 39th International Conference on Machine Learning, PMLR 162, 2022, pp. 4087-4111.
14. Yoshihiro Okawa, and Kenichi Kobayashi, "Concept Drift Detection via Boundary Shrinking", *2021 International Joint Conference on Neural Networks (IJCNN)*, Shenzhen, China, 2021, pp. 1-8, doi: 10.1109/IJCNN52387.2021.9533334.
15. Fuccellaro, M.: Concept Drift: Detection, Update and Correction. PhD Thesis, Université de Lille (2024).
16. Mahdi, Osamah A., Eric Pardede, Savitri Bevinakoppa, and Nawfal Ali. 2025. "Federated Learning Under Concept Drift: A Systematic Survey of Foundations, Innovations, and Future Research Directions" *Electronics* Vol. 14, No. 22: 4480. https://doi.org/10.3390/electronics14224480.