INFORMS Journal on Data Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Mutual Information Surprise: Rethinking Unexpectedness in Autonomous Systems

Yinsong Wang, Quan Zeng, Xiao Liu, Yu Ding

To cite this article:
Yinsong Wang, Quan Zeng, Xiao Liu, Yu Ding (2026) Mutual Information Surprise: Rethinking Unexpectedness in Autonomous Systems. INFORMS Journal on Data Science

Published online in Articles in Advance 04 Aug 2026

. https://doi.org/10.1287/ijds.2026.0182

https://pubsonline.informs.org/journal/ijds

INFORMS JOURNAL ON DATA SCIENCE
Articles in Advance, pp. 1–22
ISSN 2694-4022 (print), ISSN 2694-4030 (online)

Mutual Information Surprise: Rethinking Unexpectedness in Autonomous Systems

Yinsong Wang,[a] Quan Zeng,[a] Xiao Liu,[a] Yu Ding[a,*]

[a] H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332

*Corresponding author

ywang4542@gatech.edu (Y. Wang), qzeng47@gatech.edu (Q. Zeng), xiao.liu@isye.gatech.edu (X. Liu), yu.ding@isye.gatech.edu (Y. Ding)

https://orcid.org/0000-0001-8232-5586 (Y. Wang), https://orcid.org/0009-0003-3022-7265 (Q. Zeng), https://orcid.org/0000-0003-4510-2847 (X. Liu), https://orcid.org/0000-0001-6936-074X (Y. Ding)

Received: September 1, 2025
Revised: April 8, 2026
Accepted: June 16, 2026
Published Online in Articles in Advance: August 4, 2026

https://doi.org/10.1287/ijds.2026.0182

Abstract. A community of researchers appears to think that a machine can be surprised and have introduced various surprise measures, principally the Shannon surprise and the Bayesian surprise. The questions of what constitutes a surprise and how to react to one still elicit debates. In this work, we introduce *mutual information surprise* (MIS), a new framework that redefines surprise not as an anomaly measure, but as a signal of epistemic growth. Furthermore, we develop a statistical test sequence that could trigger a surprise reaction and propose a MIS-based reaction policy that dynamically governs system behavior through sampling adjustment and process forking. Empirical evaluations—on both synthetic domains and a dynamic pollution map estimation task—show that a system governed by the MIS-based reaction policy significantly outperforms those under classical surprise-based approaches in stability, responsiveness, and predictive accuracy. The important implication of our new proposal is that MIS quantifies the impact of new observations on mutual information, shifts surprise from reactive to reflective, enables reflection on learning progression, and thus offers a path toward self-aware and adaptive autonomous systems. We expect the new surprise measure to play a critical role in further advancing autonomous systems on their ability to learn and adapt in a complex and dynamic environment.

History: Kwok Tsui served as the senior editor for this article.
Funding: This work was supported in part by the National Science Foundation [Grant CNS-2328395].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/ijds.2026.0182.

Keywords: adaptive sampling • autonomous systems • entropy • mutual information

1. Introduction

In July 2020, *Nature* published a cover story (Burger et al. 2020) about an autonomous robotic chemist—locked in a laboratory for a week with no external communication—independently conducting experiments to search for improved photocatalysts for hydrogen production from water. In the years that followed, *Nature* featured three more articles (Merchant et al. 2023, Szymanski et al. 2023, Dai et al. 2024) highlighting the transformative role of autonomous systems in materials discovery, experimentation, and even manufacturing, each reporting orders-of-magnitude improvements in efficiency. These reports spotlighted the intensifying global race to advance autonomous technologies beyond the already well-established domain of self-driving cars (Levinson et al. 2011, MacLeod et al. 2020, Yurtsever et al. 2020, Bogdoll et al. 2022). *Nature* was not alone; numerous other outlets have documented the surge in autonomous research and innovation (Park and Tran 2012, Reis et al. 2021, Leng et al. 2023). This rapid expansion is a natural consequence of recent advances in robotics and artificial intelligence, which continue to push the boundaries of what autonomous systems can accomplish.

The systems featured in the *Nature* publications demonstrate highly capable bodies that can perform complex tasks. Recall that an autonomous system comprises two fundamental components: a brain and a body—colloquial terms for its control mechanism and its sensing-action capabilities, respectively. Unlike traditional automation systems, which follow predefined instructions to execute simple, repetitive tasks, true autonomy requires a higher level of cognitive capacity—an autonomous system is supposedly capable of making decisions with minimal human intervention. However, the brain function of the current systems, although more sophisticated than rigid preprogrammed instructions, remains relatively limited.

Surveying the literature over the past decade, we found that Nikolaev et al. (2016), Chang et al. (2020), and Burger et al. (2020) rely on classical Bayesian optimization to guide system decisions—a technique that, although effective, does not constitute full autonomy, that is, eliminating human involvement. More recent works in *Nature* (Merchant et al. 2023, Szymanski et al. 2023) continue in a similar vein, adopting active learning frameworks akin to Bayesian optimization, without fundamentally enhancing the cognitive capabilities of these systems. The conceptual limitations of their decision-making mechanisms continue to impede progress toward genuine autonomy. Ahmed et al. (2025) argue that a *core deficiency* of current autonomous systems is the absence of a "surprise" mechanism—the capacity to detect and adapt to unforeseen situations. Without this capability, true autonomy remains out of reach.

What is a "surprise," and how does it differ from existing measures governing automation? Surprise is a fundamental psychological trigger that enables humans to react to unexpected events. Intuitively, it arises when observations deviate from expectations. Traditionally, unexpectedness has been loosely equated with anomalies—quantifying inconsistencies between new observations and historical data. Common approaches to anomaly detection include statistical methods such as z-scores (Zhou and Tang 2016) and hypothesis testing (Cohen and Zhao 2015, Kamenik and Szewc 2023); distance-based techniques (Weller-Fahy et al. 2014), including Euclidean (Montechiesi et al. 2016) and Mahalanobis distances (Wang et al. 2013, Hou et al. 2020); and machine learning–based models (Schlegl et al. 2019, Lian et al. 2022), which learn patterns to identify and filter out anomalous data. However, researchers increasingly recognize that simply detecting and discarding unexpected events is insufficient for achieving higher levels of autonomy. In human cognition, unexpectedness is not inherently undesirable; in fact, surprise often signals opportunities for discovery rather than error. Although mathematically similar to anomaly measures, surprise is conceptually distinct: It is not merely a deviation to be rejected, but a valuable learning signal that can enhance adaptation and decision making.

This shift in perspective aligns with formal definitions of surprise in information theory and computational psychology, such as Shannon surprise (Barto et al. 2013), Bayesian surprise (Itti and Baldi 2009), *Bayes factor (BF) surprise* (Liakoni et al. 2021), and confidence-corrected (CC) surprise (Faraji et al. 2018). These surprise definitions quantify unexpectedness by modeling deviations from prior beliefs or probability distributions. Using some of the existing surprise definitions, Ahmed et al. (2025) demonstrated that treating surprising events not as noise to be removed but as catalysts for learning can significantly enhance a system's learning speed. Additional empirical evidence shows that incorporating surprise as a learning mechanism can improve autonomy in domains such as autonomous driving (Çatal et al. 2020, Zamiri-Jafarian and Plataniotis 2022, Dinparastdjadid et al. 2023) and manufacturing (Jin et al. 2022, Raihan et al. 2024). In Section 2, we will delve deeper into these existing measures and evaluate whether they truly serve the intended role of identifying opportunities, as human surprise does, more than merely flagging anomalies.

In this paper, we consider a general class of input-output systems described by a functional mapping, $f(\cdot)$, such that

$$f : \mathbf{x} \in \mathcal{X} \rightarrow \mathbf{y} \in \mathcal{Y}, \tag{1}$$

where $\mathbf{x}$ denotes the inputs and $\mathbf{y}$ denotes the outputs to the system. The data pair, $\{\mathbf{x}, \mathbf{y}\}$, is drawn independently (not necessarily identical) from a potentially nonstationary joint distribution $P(\mathbf{x}, \mathbf{y})$.

Most existing anomaly and surprise measures implicitly assume that the underlying distribution $P(\mathbf{x}, \mathbf{y})$ remains stationary over time. Such an assumption is restrictive: truly autonomous systems operate in evolving environments, where system dynamics, external conditions, and even objectives may shift. A surprise measure should not merely detect deviations under a fixed model, but rather actively capture potential changes in both inputs and the systems.

Motivated by the need to address this limitation, we argue that it is essential to develop a new surprise measure that does not anchor itself to the stationarity of $P(\mathbf{x}, \mathbf{y})$, but instead inherently fosters learning and deepens an autonomous system's understanding of the underlying processes it encounters. To capture this dynamic capability, we introduce the mutual information surprise (MIS)—a new framework that redefines how autonomous systems interpret and respond to unexpected events. MIS quantifies the degree of both *frustration* and *enlightenment* associated with new observations, measuring their impact on refining the system's internal understanding of its environment. We contrast the outcomes made by applying mutual information surprise with that by relying solely on classical surprise definitions to highlight MIS's potential to enhance learning and decision making.

The paper is organized as follows. In Section 2, we revisit the concept of surprise by presenting a taxonomy of existing surprise measures and introducing the intuition, mathematical formulation, and limitations of classical definitions. In Section 3, we formally define the MIS and derive a testing sequence for detecting multiple types of system changes in autonomous systems. We also design an MIS reaction policy (MISRP) that provides high-level

guidance to complement existing exploration-exploitation active learning strategies. In Section 4, we compare MIS with classical surprise measures to illustrate its numerical stability and enhanced cognitive capability. We further demonstrate the effectiveness of the MIS reaction policy through a pollution map estimation simulation. In Section 5, we conclude the paper.

2. Current Surprise Definitions and Their Limitations

Classical definitions of surprise, such as Shannon and Bayesian surprise, provide elegant mathematical frameworks for quantifying unexpectedness. Indeed, the notion of surprise plays a central role in the field of active inference (Friston et al. 2015), where it serves as a key driver of perception and action. However, these approaches often fall short of capturing the mechanisms that underlie adaptive behavior, namely continuous learning and flexible model updating. In this section, we revisit existing formulations of surprise, examining their conceptual foundations while highlighting both their strengths and their limitations.

Before proceeding with our discussion, we introduce the notation used throughout this paper. Scalars are denoted by lowercase letters (e.g., x), vectors by bold lowercase letters (e.g., $\mathbf{x}$), and matrices by bold uppercase letters (e.g., $\mathbf{X}$). Distributions in the data space are represented by uppercase letters (e.g., P), probabilities by lowercase letters (e.g., p), and distributions in the parameter space by the symbol π. The L_2 norm is denoted by $\|\cdot\|_2$, and the absolute value or L_1 norm is denoted by $|\cdot|$. We use $\mathbb{E}[\cdot]$ to denote the expectation operator and $\mathrm{sgn}(\cdot)$ for the sign operator. Estimators are denoted with a hat, as in $\hat{\cdot}$.

2.1. Family of Shannon Surprises

The family of Shannon surprise metrics emphasizes the improbability of observed data, typically *independent* of explicit model parameters. This class broadly aligns with "observation" and "probabilistic-mismatch" surprises as categorized in Modirshanechi et al. (2022). The central question which the Shannon family of surprises tries to answer is: How unlikely is the observation? For that, a Shannon surprise (Baldi 2002) is commonly defined as

$$S_{\text{Shannon}}(\mathbf{x}) = -\log p(\mathbf{x}), \tag{2}$$

interpreting surprise directly through event rarity. The above expression defines Shannon surprise in terms of the marginal input observation $\mathbf{x}$. It is not difficult to see that the same definition is applicable to outputs $\mathbf{y}$. For a practical system with an input-output structure as expressed in Equation (1), surprise is more naturally defined with respect to the predictive distribution of outputs given inputs, namely in a conditional perspective. This is to say, instead of evaluating $-\log p(\mathbf{x})$, one considers

$$S_{\text{Shannon}}(\mathbf{x}, \mathbf{y}) = -\log p(\mathbf{y}|\mathbf{x}) \tag{3}$$

for Shannon surprise. In this work, we adopt this conditional viewpoint when evaluating the Shannon family of surprise measures.

Although conceptually clear, the above definition has a significant limitation: Encountering a Shannon surprise does not inherently imply knowledge acquisition. Consider, for instance, a uniform dartboard—a stochastic yet entirely understood system. Each outcome has an equally low probability, thus appearing "surprising" under Shannon's definition, despite humans neither genuinely finding these outcomes surprising nor gaining any additional knowledge by observing them. In other words, the focus of Shannon surprise is statistical rarity rather than knowledge gain.

To address this limitation, particularly in highly stochastic scenarios, *residual information surprise* (Dinparastdjadid et al. 2023) has been introduced, which measures surprise by quantifying the gap between the minimally achievable and observed Shannon surprises:

$$S_{\text{Residual}}(\mathbf{x}, \mathbf{y}) = \left|\min_{\mathbf{y}'}\{-\log p(\mathbf{y}'|\mathbf{x})\} - (-\log p(\mathbf{y}|\mathbf{x}))\right| = \max_{\mathbf{y}'} \log p(\mathbf{y}'|\mathbf{x}) - \log p(\mathbf{y}|\mathbf{x}).$$

In the dartboard example, residual information surprise becomes zero for all outcomes, because $p(\mathbf{y}'|\mathbf{x})$ remains constant for every outcome $\mathbf{y}'$, accurately reflecting an absence of genuine surprise. However, this formulation introduces a conceptual challenge, as determining $\max_{\mathbf{y}'} \log p(\mathbf{y}'|\mathbf{x})$ implicitly presumes an omniscient oracle, an assumption typically infeasible in practice.

Interestingly, Shannon surprise serves as a foundation for various anomaly measures. For example, under Gaussian assumptions, Shannon surprise becomes proportional to squared error:

$$S_{\text{Shannon}}(\mathbf{x}, \mathbf{y}) \propto \|\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}|\mathbf{x}}\|_2^2,$$

thus linking surprise with deviation from the mean. Similarly, assuming a Laplace distribution, Shannon

surprise recovers an absolute error interpretation, termed *absolute error surprise* in Prat-Carrabin et al. (2021):

$$S_{\text{Shannon}}(\mathbf{x}, \mathbf{y}) \propto |\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}|\mathbf{x}}|.$$

Both squared error surprise and absolute error surprise are commonly utilized metrics in anomaly detection literature (Rousseeuw and Croux 1993, Aytekin et al. 2018, Nguyen et al. 2019).

2.2. Family of Bayesian Surprises

Bayesian surprises, by contrast, explicitly model belief updates. Rather than assessing the improbability of an observation, the Bayesian family assumes that the data-generating process is governed by an underlying parameter $\boldsymbol{\theta}$ and defines surprise in terms of how a new observation changes the posterior belief over $\boldsymbol{\theta}$. These measures quantify the degree to which a new observation alters the internal model, shifting the focus from event rarity to epistemic impact. This concept parallels the "belief-mismatch" surprise in the taxonomy by Modirshanechi et al. (2022).

The canonical formulation, introduced in Baldi (2002), defines Bayesian surprise as the Kullback-Leibler (KL) divergence between the prior and posterior distributions over parameters:

$$S_{\text{Bayes}}(\mathbf{x}) = D_{\text{KL}}(\pi(\boldsymbol{\theta}|\mathbf{x}) \| \pi(\boldsymbol{\theta})). \tag{4}$$

The above definition is presented in terms of $\mathbf{x}$. Like in the case of Shannon surprise, the Bayesian family of surprise measures naturally extend to input-output systems as well. This extension follows from recognizing that the parameter $\boldsymbol{\theta}$ characterizes the functional relationship between the input $\mathbf{x}$ and the output $\mathbf{y}$. As such, observations arrive as paired data $(\mathbf{x}, \mathbf{y})$, and belief updates occur over the parameters governing the conditional mapping $P(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta})$. Consequently, the Bayesian Surprise for input-output systems can be written as

$$S_{\text{Bayes}}(\mathbf{x}, \mathbf{y}) = D_{\text{KL}}(\pi(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y}) \| \pi(\boldsymbol{\theta})), \tag{5}$$

where the posterior $\pi(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})$ reflects the updated belief after observing the input-output pair.

The Bayesian Surprise measure offers a principled approach to belief revision and naturally aligns with learning mechanisms. In theory, it encourages agents to reduce surprise through model updates, providing a pathway toward adaptive autonomy.

However, Bayesian surprise is not without limitations. As data accumulate, new observations exert diminishing influence on the posterior, rendering the agent increasingly "stubborn." This behavior can result in Bayesian surprise overlooking rare but meaningful anomalies. For example, consider the discovery by S. S. Ting of the J particle, characterized by an unusually long lifespan compared with other particles in its class. Under standard Bayesian updating, scientists' beliefs about particle lifespans would barely shift due to this single observation. Consequently, Bayesian surprise would classify such an event as merely an anomaly, potentially disregarding it.

To mitigate this posterior overconfidence, the confidence-corrected (CC) surprise (Faraji et al. 2018) compares the current informed belief against that of a naïve learner with a flat prior:

$$S_{\text{CC}}(\mathbf{x}, \mathbf{y}) = D_{\text{KL}}\left(\pi(\boldsymbol{\theta}) \| \pi'(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})\right),$$

where $\pi'(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y})$ represents the updated belief assuming a uniform prior. This confidence-corrected formulation remains sensitive to new data irrespective of prior history. In the J particle example, employing CC surprise would trigger a genuine surprise, because the posterior remains responsive to the novel observation without the inertia introduced by extensive historical data.

A related idea emerges with the Bayes Factor (BF) surprise (Liakoni et al. 2021), which compares likelihoods under naïve and informed beliefs:

$$S_{\text{BF}}(\mathbf{x}, \mathbf{y}) = \frac{p(\mathbf{x}, \mathbf{y}|\pi^0(\boldsymbol{\theta}))}{p(\mathbf{x}, \mathbf{y}|\pi^t(\boldsymbol{\theta}))},$$

where $\pi^0(\boldsymbol{\theta})$ represents the naïve (untrained) prior and $\pi^t(\boldsymbol{\theta})$ the informed belief based on all prior observations up to time t (before observing $(\mathbf{x}, \mathbf{y})$). This ratio quantifies how strongly the current observation supports the naïve prior over the informed prior. In practice, the effectiveness of both CC and BF surprises heavily depends on constructing appropriate priors—a task often challenging and subjective.

Another variant within the Bayesian Surprise family is *Postdictive Surprise* (Kolossa et al. 2015), which operates in the output space rather than parameter space as in the original Bayesian surprise:

$$S_{\text{Postdictive}}(\mathbf{x}, \mathbf{y}) = D_{\text{KL}}(P(\mathbf{y}|\boldsymbol{\theta}', \mathbf{x}) \| P(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x})), \tag{6}$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ denote the predictive model parameters before and after the update, respectively. Kolossa et al.

(2015) argue that computing KL divergence in the output space is more computationally tractable for variational models but potentially less expressive when output variance depends on the input (e.g., under heteroskedastic conditions).

2.3. Reflection

We acknowledge the presence of alternative categorizations of surprise definitions, notably the taxonomy in Modirshanechi et al. (2022), which classifies surprise measures into three groups: observation surprises, probabilistic-mismatch surprises, and belief-mismatch surprises. As discussed previously, the Shannon surprise family aligns closely with the first two categories, whereas the Bayesian surprise family corresponds to the last.

These categorizations are not strictly delineated. For instance, residual information surprise incorporates a conceptual element common to the Bayesian surprise family—providing a baseline against which the observed data are contrasted with. On the other hand, BF surprise, despite being explicitly Bayesian in its formulation, closely resembles a Shannon surprise conditioned on alternative priors. Notwithstanding their philosophical distinctions, Bayesian and Shannon surprises often behave similarly in practice; we provide further details on this observation in Section 4.

It is understandable that researchers initially explored these two foundational surprise definitions, each possessing inherent limitations: Shannon surprise conflates probability with knowledge gain, whereas Bayesian surprise suffers from increasing posterior stubbornness. Subsequent refinements emerged to address these shortcomings, primarily through adjusting the choice of prior to create more meaningful contrasts. The residual information surprise assumes an oracle-like prior, whereas CC and BF surprises rely on a noninformative prior. Regardless of the priors chosen, defining a suitable prior remains a challenging and unresolved issue in the research community.

Both surprise families share other critical limitations: They are *single-instance measures* by design and are inherently *one-sided*. Being single instance means that they primarily focus on surprise based solely on the marginal impact of individual observations, without explicitly modeling cumulative learning dynamics over time, whereas being one-sided means that they have a decision threshold on one single side, offering limited expressiveness because human perceptions of surprise can range from positive to negative.

3. MIS

In this section, we introduce the concept of MIS. We first explore the intuition and motivation underlying this concept, followed by the development of a novel, theoretically grounded testing sequence. We then discuss the implications when this test sequence is violated and propose a reaction policy contingent on different types of violations.

3.1. Alternative Surprise Definition

Recall that we consider input-output systems described by a functional mapping $f(\cdot): \mathbf{x} \to \mathbf{y}$, with observations drawn from a joint distribution $P(\mathbf{x}, \mathbf{y})$. Classical surprise measures are implicitly constructed under the assumption that $P(\mathbf{x}, \mathbf{y})$ remains stationary over time. However, autonomous engineering systems often operate in evolving environments where such stationarity assumptions may not hold. What this means is that not all changes in $P(\mathbf{x}, \mathbf{y})$ are of interest or should trigger a surprise reaction. This prompts us to look for a new tool to define surprise for the input-output system we are dealing with.

Our research leads us to the concept of *mutual information* (MI), which was introduced by Shannon (1948) (although the specific term, "mutual information," was coined by Robert Fano nearly 10 years later) and defined as

$$I(\mathbf{x};\mathbf{y}) = \mathbb{E}_{\mathbf{x},\mathbf{y}}\left[\log\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y})}\right] = H(\mathbf{x}) + H(\mathbf{y}) - H(\mathbf{x},\mathbf{y}) = H(\mathbf{y}) - H(\mathbf{y}|\mathbf{x}), \tag{7}$$

where $H(\cdot)$ denotes entropy, which measures the uncertainty of a random variable. The mutual information, $I(\mathbf{x};\mathbf{y})$, which is always nonnegative, quantifies the reduction in uncertainty about $\mathbf{y}$ when $\mathbf{x}$ is observed. In fact, mutual information can serve as a quantitative tool to benchmark "system learning" because Fano's inequality (Verdú 1994) establishes that independently of prediction models, the best possible learning performance is governed by mutual information. Specifically, an increasing $I(\mathbf{x};\mathbf{y})$ indicates that the system is gaining in understanding of the underlying functional relationship $f(\cdot)$, whereas stagnation or decline in $I(\mathbf{x};\mathbf{y})$ suggests that the learning has stalled. Moreover, not all changes in $P(\mathbf{x}, \mathbf{y})$ will cause a change in mutual information, a property that would be useful to capture genuine surprises in dynamic autonomous systems.

Unlike the traditional change detection literature which usually defines the null hypothesis (the baseline) as a stationary in-control distribution, we define our null hypothesis as an input-output system with a stationary mutual information. We formalize this null hypothesis through the following notion of a well-regulated autonomous system.

Definition 1. A well-regulated input-output autonomous system may exhibit nonstationary joint distributions $P(\mathbf{x}, \mathbf{y})$ over time, yet its mutual information $I(\mathbf{x}; \mathbf{y})$ remains unchanged.

For an autonomous system to trigger a surprise, it is when the mutual information of the system changes, meaning that the system is out of the well-regulated region it has been operating under. In practice, the mutual information $I(\mathbf{x}; \mathbf{y})$ is typically estimated using maximum likelihood estimation (MLE) (Paninski 2003); some details regarding the estimator used in this work are provided in the Online Appendix. The estimation introduces uncertainty, so that statistical decision thresholds need to be established as a function of the prescribed significance level; we will provide the upper/lower statistical decision thresholds in Section 3.2.

The above arguments lead us to introduce a MIS, defined as the change in estimated mutual information after incorporating new observations:

$$\text{MIS} \triangleq \hat{I}_{n+m} - \hat{I}_n, \tag{8}$$

where $\hat{I}_n$ denotes the estimate of mutual information based on the first n observations, and $\hat{I}_{n+m}$ denotes the estimate after observing an additional m samples, with $\mathbf{x}$ and $\mathbf{y}$ omitted for simplicity. In this work, we adopt the standard estimation approach of histogram-based distribution representation (Paninski 2003) and a default of 10 bins unless otherwise specified.

A large positive MIS indicates *enlightenment* (increase in $I(\mathbf{x}; \mathbf{y})$), meaning that the new observations significantly improve the system understanding of the input-output relationship. In contrast, a near-zero or negative MIS indicates *frustration* (decrease in $I(\mathbf{x}; \mathbf{y})$), suggesting that new data contributes little to learning or even disrupts previously acquired structure. In this way, MIS provides a practical signal for assessing whether an autonomous system continues to acquire knowledge from its environment. We argue that MIS is a better metric for guiding decision making of autonomous systems, and we will provide further support to substantiate our claim in the subsequent sections.

3.2. Bounding MIS

Testing the change in $I(\mathbf{x}; \mathbf{y})$ via MIS is challenging: Mutual information estimation is nonlinear and exhibits complex variance. The standard method, although principled, is a computationally expensive permutation test (François et al. 2006, Doquire and Verleysen 2013), involving repeatedly shuffling $m + n$ observations into two groups, calculating MI differences, and evaluating rejection probabilities:

$$p = \frac{1}{B} \sum_{i=1}^{B} \mathbf{1}(|\Delta \hat{I}| > |\Delta \hat{I}|_i),$$

where $\Delta \hat{I} = \hat{I}_n - \hat{I}_m$ represents the actual differences between mutual information estimations, and $\Delta \hat{I}_i$ represents the ith permuted difference; $\mathbf{1}(\cdot)$ is the indicator function. In real-time streaming scenarios, however, permutation tests become impractical due to their computational load. Moreover, when $m \ll n$, permutation tests lose effectiveness, yielding noisy outcomes.

An alternative is standard deviation–based testing. For MLE mutual information estimator $\hat{I}_n$, its estimation standard deviation satisfies (Paninski 2003)

$$\sigma \lesssim \frac{\log n}{\sqrt{n}}, \tag{9}$$

where $\lesssim$ stands for less or equal to (in terms of order), which yields an analytical test on the mutual information change when omitting the bias term (brief derivation provided in the Online Appendix):

$$\hat{I}_{m+n} - \hat{I}_n \in \pm \sqrt{\frac{\log^2(m+n)}{m+n} + \frac{\log^2 n}{n}} \cdot z_\alpha \asymp \mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right), \tag{10}$$

where z_α represents the standard normal random variable at confidence level α and $\asymp$ represents equal in order. But this test too is unsatisfying, because the above bound is so loose that it rarely gets violated. The root cause is the loose upper bound shown in Equation (9), where empirical evidence suggests the true estimation standard

deviation is usually much smaller than the theoretical bound. We provide the empirical evidence in the Online Appendix.

Therefore, we turn to a new path for bounding MIS as follows. First, we impose several mild assumptions on the observations and the physical process.

Assumption 1. *We impose the following assumptions on the sampling process and physical system.*

1. *We assume that the existing observations are typical in the sense of the asymptotic equipartition property (Cover 1999), meaning that empirical statistics computed from the data are representative of their corresponding expected values under the experimental design's intended distribution, that is, $\hat{I}_n \approx \mathbb{E}[\hat{I}_n]$. This is true when we regard the initial observations as true system information.*
2. *The number of existing observations n is much smaller than the cardinality of space $\mathcal{X}, \mathcal{Y}$. $n \ll |\mathcal{X}|, |\mathcal{Y}|$*
3. *The number of new observations m is much smaller than the number of existing observations. $m \ll n$.*

Theorem 1. *Consider a well-regulated autonomous system defined in Definition 1, which satisfies the conditions in Assumption 1. With probability at least $1-\rho$, the change in MLE-based mutual information estimates satisfies*

$$\hat{I}_{n+m} - \hat{I}_n \in (\log(m+n) - \log n) \pm \frac{\sqrt{2m \log \frac{2}{\rho}} \log (m+n)}{m+n} \triangleq MIS_{\pm}.$$

$MIS_{\pm}$ denotes the upper and lower bound for the test sequence.

The proof of Theorem 1 is shown in the Online Appendix. These bounds are both tighter $\left(\mathcal{O}\left(\frac{\log n}{n}\right)\right.$ instead of $\left.\mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right)\right)$ and more efficient (analytical test sequence) than previous methods. The bounds offer theoretically grounded thresholds within which we expect MI to evolve. When these bounds $MIS_{\pm}$ are breached—either from below or from above—we then deem that the true mutual information of the system has changed.

3.3. What Does MIS Actually Tell Us?

When the quantity MIS = $\hat{I}_{n+m} - \hat{I}_n$ falls outside the established bounds $MIS_{\pm}$—either exceeding the upper bound or falling below the lower bound—the system is considered surprised, thereby triggering an MIS. Essentially, Theorem 1 functions as a statistical hypothesis test: The null hypothesis posits that the underlying system remains well regulated, implying $\Delta I = I_{n+m} - I_n = 0$, where I_n denotes the true mutual information at the time of n observations. Any violation indicates a significant shift, with negative deviations ($\Delta I < 0$) and positive deviations ($\Delta I > 0$) each carrying distinct implications on learning performance via the Fano lemma (Verdú 1994).

Recall that mutual information can be expressed in terms of entropy, as shown in Equation (7), so changes in ΔI may result from variations in $H(\mathbf{x})$, $H(\mathbf{y})$, and $H(\mathbf{y}|\mathbf{x})$. In this section, we examine the implications of MIS under different driving forces.

3.3.1. Violation from Below: Learning Has Stalled or Regressed.

If

$$\text{MIS} < \text{MIS}_{-},$$

this implies $\Delta I(\mathbf{x};\mathbf{y}) < 0$, signifying a downward shift in mutual information. A negative surprise indicates diminished or stalled learning, potentially due to the following:

1. **Stagnation in Exploration:** A downward shift driven by a decrease in input entropy $\Delta H(\mathbf{x}) < 0$ suggests the system repeatedly samples in a limited region, thus gathering redundant data with minimal new information.
2. **Increased Noise or Process Drift:** A downward shift could also result from increased conditional entropy $\Delta H(\mathbf{y}|\mathbf{x}) > 0$, indicating greater uncertainty in predicting $\mathbf{y}$ given $\mathbf{x}$. Practically, this often signifies increased external noise or a fundamental change in the underlying process.

3.3.2. Violation from Above: Sudden Growth in Understanding.

If

$$\text{MIS} > \text{MIS}_{+},$$

this implies $\Delta I(\mathbf{x};\mathbf{y}) > 0$, indicating an upward shift in mutual information. This positive surprise can result from the following:

1. **Aggressive Exploration:** If the increase is driven by higher input entropy $\Delta H(\mathbf{x}) > 0$, the system is likely exploring previously unvisited regions aggressively, potentially inflating knowledge gains without sufficient validation.

2. **Reduction in Noise:** An increase due to reduced conditional entropy $\Delta H(\mathbf{y}|\mathbf{x}) < 0$ signals a desirable decrease in uncertainty, thus generally representing a beneficial development.
3. **Novel Discovery:** An increase in output entropy $\Delta H(\mathbf{y}) > 0$ suggests discovery of novel and previously rare outputs—particularly valuable in exploratory or scientific contexts.

Table 1 summarizes potential causes for MIS violations and their implications. These patterns help the system differentiate between meaningful learning and misleading deviations, expanding beyond the capacity of classical surprise measures and providing a road map to corrective or adaptive responses for higher level autonomy. We purposely omit the case where a decrease in $H(\mathbf{y})$ causes violation from below, because this scenario typically lacks independent significance. Instead, its happening is generally caused by changes in sampling strategy or underlying processes, which we have already discussed.

3.4. Reaction Policy: Three-Pronged Approach

Following the identification of potential causes behind MIS triggers (Section 3.3), the next question is how the system should respond. Naturally, the system's reaction should align with the dominant entropy component contributing to the change. In practice, we identify the dominant entropy change by computing and ranking the ratios

$$\frac{\text{sgn(MIS)}\Delta\hat{H}(\mathbf{x})}{|\text{MIS}|}, \quad \frac{\text{sgn(MIS)}\Delta\hat{H}(\mathbf{y})}{|\text{MIS}|}, \quad \text{and} \quad \frac{\text{sgn(MIS)}\Delta\hat{H}(\mathbf{y}|\mathbf{x})}{|\text{MIS}|},$$

where $\Delta\hat{H}(\cdot) = \hat{H}_{m+n}(\cdot) - \hat{H}_n(\cdot)$ denotes the estimated entropy change.

We do not prescribe a specific reaction when $\Delta\hat{H}(\mathbf{y})$ dominates the MIS, as an increase in $H(\mathbf{y})$ is typically a passive consequence of changes in $H(\mathbf{x})$ and $H(\mathbf{y}|\mathbf{x})$. When both $H(\mathbf{x})$ and $H(\mathbf{y}|\mathbf{x})$ remain relatively stable, a rise in $H(\mathbf{y})$ indicates that the current sampling strategy is effectively uncovering novel information; thus, no change in action is required.

For $\Delta\hat{H}(\mathbf{x})$ and $\Delta\hat{H}(\mathbf{y}|\mathbf{x})$, situations may arise where their contributions are similar, that is, no clear dominant entropy component exists, and we need a resolution mechanism to break the tie. To address all these scenarios, we propose a three-pronged reaction policy that serves as a supervisory layer, compatible with existing exploration-exploitation sampling strategies.

1. **Sampling Adjustment.** The first policy addresses variations in input entropy $H(\mathbf{x})$. If $\Delta\hat{H}(\mathbf{x}) > 0$ dominates MIS, indicating overly aggressive exploration, the system should moderate exploration and emphasize exploitation to prevent fitting to noise. Conversely, if $\Delta\hat{H}(\mathbf{x}) < 0$, suggesting redundant sampling, the system should enhance exploration to restore sample diversity.
2. **Process Forking.** The second policy responds to variations in conditional entropy $H(\mathbf{y}|\mathbf{x})$, that is, changes in conditional distribution. Upon surprise triggered by $\Delta\hat{H}(\mathbf{y}|\mathbf{x})$, the system forks into two subprocesses, each consisting of n existing observations and m new observations divided at the surprise moment (Theorem 1). The two subprocesses represent the prior process (existing observations) and the likely altered process (new observations), and will continue their sampling separately. The subprocess first encountering a $\Delta\hat{H}(\mathbf{y}|\mathbf{x})$-triggered surprise is discarded, and the remaining subprocess continues as the main process. In the extremely rare case when both subprocesses trigger a $\Delta\hat{H}(\mathbf{y}|\mathbf{x})$ dominated MIS surprise at the same time, we discard the process with fewer observations and continue with the subprocess with more observations.
3. **Coin Toss Resolution.** There are occasions where changes in $\Delta\hat{H}(\mathbf{x})$ and $\Delta\hat{H}(\mathbf{y}|\mathbf{x})$ are comparable, making selecting a reaction policy challenging. Instead of arbitrarily favoring the slightly larger change, we always use a biased coin toss approach, stochastically selecting which entropy to address based on the magnitude of changes:

$$p_{\text{adjust}} = \frac{|\Delta\hat{H}(\mathbf{x})|}{|\Delta\hat{H}(\mathbf{x})| + |\Delta\hat{H}(\mathbf{y}|\mathbf{x})|}, \quad p_{\text{fork}} = 1 - p_{\text{adjust}}.$$

Table 1. Summary Table: MIS Violations and Their Potential Causes

Violation type	Possible causes	Trend in mutual information
Violation from below	Stagnation in exploration	$\downarrow H(\mathbf{x}) \Rightarrow \downarrow I(\mathbf{x};\mathbf{y})$
	Increased noise/process drift	$\uparrow H(\mathbf{y}\|\mathbf{x}) \Rightarrow \downarrow I(\mathbf{x};\mathbf{y})$
Violation from above	Aggressive exploration	$\uparrow H(\mathbf{x}) \Rightarrow \uparrow I(\mathbf{x};\mathbf{y})$
	Noise reduction	$\downarrow H(\mathbf{y}\|\mathbf{x}) \Rightarrow \uparrow I(\mathbf{x};\mathbf{y})$
	Novel discovery	$\uparrow H(\mathbf{y}) \Rightarrow \uparrow I(\mathbf{x};\mathbf{y})$

The decision variable z is sampled as $z \sim \text{Bernoulli}(P_{\text{adjust}})$, with $z = 1$ indicating sampling adjustment and $z = 0$ indicating process forking. This mechanism ensures balanced reactions and robustness and prevents overreactions to marginal signals.

The description above provides a brief summary of the MIS reaction policy. In the remaining portion of this section, we will present the specific MIS reaction policy in an algorithm. To do that, we first need to define a sampling process formally and then present the detailed algorithmic implementation of this reaction policy in Algorithm 1.

Definition 2. A sampling process $\mathcal{P}(\mathbf{X}, g(\cdot))$ consists of two components: existing observations $\mathbf{X}$ and a sampling function $g(\cdot)$, where the next sample location is determined by

$$\mathbf{x}_{\text{next}} \sim g(\mathbf{X}),$$

with $\mathbf{x}_{\text{next}}$ drawn from the stochastic oracle $g(\mathbf{X})$. If $g(\cdot)$ is deterministic, $\sim$ is replaced by equality (=). For clarity, a sampling process with n existing observations is denoted $\mathcal{P}_n$.

Algorithm 1 (MISRP)

Require: A sampling process $\mathcal{P}(\mathbf{Z}, g(\cdot))$, where $\mathbf{Z}$ consists of k pairs of input $\mathbf{X}$ and output $\mathbf{Y}$; A maximum reflection threshold T; Reflection period $m = 2$

while $m \le \min(T, \frac{k}{2})$ **do**
  Set $n = k - m$; Compute $MIS = \hat{I}_{m+n} - \hat{I}_n$; Record $\Delta\hat{H}(\mathbf{x})$, $\Delta\hat{H}(\mathbf{y})$, and $\Delta\hat{H}(\mathbf{y}|\mathbf{x})$
  if $MIS \notin MIS_{\pm}$ and $\frac{\text{sgn(MIS)}\Delta\hat{H}(\mathbf{y})}{|\text{MIS}|} \neq \max\left\{\frac{\text{sgn(MIS)}\Delta\hat{H}(\mathbf{x})}{|\text{MIS}|}, \frac{\text{sgn(MIS)}\Delta\hat{H}(\mathbf{y})}{|\text{MIS}|}, \frac{\text{sgn(MIS)}\Delta\hat{H}(\mathbf{y}|\mathbf{x})}{|\text{MIS}|}\right\}$ **then**
    Compute bias: $p \leftarrow \frac{|\Delta\hat{H}(\mathbf{x})|}{|\Delta\hat{H}(\mathbf{x})| + |\Delta\hat{H}(\mathbf{y}|\mathbf{x})|}$
    Sample $z \sim \text{Bernoulli}(p)$
    if $z = 1$ **then** ▷ Sampling Adjustment
      if $MIS > MIS_{+}$ **then**
        Modify g to reduce exploration and increase exploitation
      else
        Modify g to increase exploration and reduce redundancy
      end if
      break while
    else ▷ Process Forking
      if $\mathcal{P}$ is forked and the other process is not requesting Process Forking **then**
        Delete $\mathcal{P}$; Merge the other process as the main process
        break while
      end if
      if $\mathcal{P}$ is forked and the other process is requesting Process Forking **then**
        Delete the $\mathcal{P}$ with fewer data; Merge the other one as the main process
        break while
      end if
      Fork process into two branches: $\mathcal{P}_n$ and $\mathcal{P}_m$
      Call MISRP($\mathcal{P}_n, t$) and MISRP($\mathcal{P}_m, t$)
      break while
    end if
  else
    No action required (surprise within expected bounds)
  end if
  $m = m + 1$
end while

We offer several remarks on the MIS reaction policy MISRP($\mathcal{P}, t$):

- In the pseudocode, we introduce two additional notations: the maximum reflection threshold T and the total number of observations k. In practice, MIS is computed retroactively; that is, given a sequence of k observations, we partition them into m recent observations and $n = k - m$ older observations to compute the MIS. We term the m recent observation as the reflection period, and we increment m to iterate over different partition points. The reflection period m is constrained to be no greater than $\min(T, \frac{k}{2})$. This constraint is motivated by the comparative behavior of test statistics derived from Theorem 1 and the variance-based test in Equation (10). Specifically, when $m = n$,

both our proposed test and the variance-based test yield statistics of order $\mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right)$. As discussed in Section 3.2, such statistics are typically too loose to be violated in practice, thereby diminishing the sensitivity advantage of our method. Consequently, evaluating MIS beyond $m = \frac{k}{2}$ is unnecessary and computationally inefficient. The reflection threshold T is introduced to ensure computational feasibility, and we recommend selecting T as large as computational resources permit.

- Note that the reflection period m starts at 2. This implies that the reaction policy does not respond to a single-instance surprise. Mathematically, this is because the derivation of the bound in Theorem 1 is ill defined for $m = 1$. Intuitively, MIS measures the progression of learning in a sampling process, and it is impossible to determine whether a single observation is informative or erroneous without additional verification. Therefore, the MIS policy always takes at least two additional samples to start to react. One may argue that this requirement for an extra sample imposes additional cost in conducting experiments. That is true. But recall one insight from the study in Ahmed et al. (2025) is the benefit of "the extra resources spent on deciding the nature of an observation" in the long run.
- It is important to emphasize that both the sampling adjustment and process forking approaches are rooted in the active learning literature and practice. Balancing exploration and exploitation, that is, sampling adjustment, has long been a key topic in Bayesian optimization and active learning (Bondu et al. 2010), whereas discarding irrelevant observations, as we do in process forking, is a common practice in the data set drift literature (Sugiyama et al. 2007, Bickel et al. 2009, Moreno-Torres et al. 2012, Žliobaitė et al. 2016, Zhang et al. 2023). Our MIS reaction framework provides a principled mechanism for autonomous systems to determine how to balance exploration versus exploitation and when or what to discard (i.e., forget).

4. Numerical Analysis

In this section, we illustrate the merits of MIS. Section 4.1 demonstrates the strength of MIS compared with classical surprise measures. Section 4.2 showcases the advantages of the MIS reaction policy in the context of dynamically estimating a pollution map using data generated from a physics-based simulator.

4.1. Putting Surprise to the Test

To compare MIS with classical surprise measures—principally Shannon and Bayesian surprises—we conduct a series of controlled simulations using a simple yet interpretable system, designed to reveal how each measure behaves under varying conditions. The system is governed by the mapping

$$y = x \mod 10, \tag{11}$$

chosen for its simplicity, modifiability, and clarity of interpretation. The first four scenarios are fully deterministic, whereas the final two introduce noise and perturbations, enabling an assessment of whether each surprise measure responds meaningfully to new observations, structural changes, or stochastic disturbances. Each simulation begins with 100 samples drawn uniformly from $x \in [0, 30]$ to establish the system's initial knowledge. We then progressively introduce new data under different conditions, recording the response of each surprise measure. As the magnitudes of MIS, Shannon surprise, and Bayesian surprise differ in scale, our analysis focuses on *behavioral trends*—how each measure changes, spikes, or saturates—rather than on their absolute values.

The surprise measures are computed as follows. Shannon surprise is calculated using its classical definition in Equation (3), namely as the negative log-likelihood of the true label under a Gaussian process (GP) predictive model. Bayesian surprise is computed using postdictive surprise, defined in Equation (6), by evaluating the KL divergence between the prior and posterior predictive distributions of $\mathbf{y}$ at each input $\mathbf{x}$.

Although the current definition of Shannon and Bayesian surprises are single instance, they can be extended to multi-instance settings. Specifically, for a sequence of m independently sampled observation pairs $(\mathbf{X}, \mathbf{Y}) = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{m}$, the cumulative Shannon surprise (CSS) is defined as

$$S_{\text{Shannon}}(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{m} S_{\text{Shannon}}(\mathbf{x}_i, \mathbf{y}_i | p_{i-1}), \tag{12}$$

where p_{i-1} denotes the predictive distribution $p(\mathbf{y}|\mathbf{x})$ prior to observing $(\mathbf{x}_i, \mathbf{y}_i)$. This formulation is also known as the *cumulative prequential log score* (Dawid 1984), which has been widely used in sequential concept drift and anomaly detection literature (Ayed et al. 2020, Lee et al. 2020, Bayram et al. 2022).

For Bayesian (postdictive) surprise, we define the multi-instance version as the divergence between predictive distributions evaluated on the most recent observation $(\mathbf{x}_m, \mathbf{y}_m)$, after updating parameters from $\boldsymbol{\theta}$ to $\boldsymbol{\theta}'$ using the

data set $(\mathbf{X}, \mathbf{Y})$. Formally,

$$S_{\text{Postdictive}}(\mathbf{X}, \mathbf{Y}) = D_{\text{KL}}(P(\mathbf{y}_m|\boldsymbol{\theta}', \mathbf{x}_m) \| P(\mathbf{y}_m|\boldsymbol{\theta}, \mathbf{x}_m)).$$

We consider two practical multi-instance variants: a *cumulative* version that aggregates surprise values from the beginning of the sequence and a *rolling-window* version that measures cumulative surprise within the most recent $W = 20$ sampling steps.

All versions of Shannon and Bayesian surprises use the same Gaussian process predictive model, with a Matérn kernel ($\nu = 2.5$) and a fixed noise level of 0.1. Except for the multi-instance Bayesian surprise, the model is retrained using all available observations after each surprise computation.

For MIS, we treat the initial 100 observations as the initial sample size $n = 100$, as defined in Section 3.1. As sampling continues, the number of new observations m increases (represented in the ticks of the x-axis in the figures). The output space has cardinality $|\mathcal{Y}| = 10$, corresponding to the 10 possible outcomes of the modulus function, except in Scenario 6, where $|\mathcal{Y}| = 20$. MIS is calculated as defined in Equation (8). When the theoretical bound in Theorem 1 is used, the probability level is set to $\rho = 0.1$. The bias term is adjusted as discussed in Section 3.2, because $n \gg |\mathcal{Y}|$ in this setting.

4.1.1. Scenario 1: Standard Exploration. New data are randomly sampled from $x \in [30, 100]$, expanding the domain while adhering to the same underlying response function in Equation (11). As the sampling region broadens, the input distribution $p(\mathbf{x})$ evolves, which in turn alters the joint distribution $p(\mathbf{x}, \mathbf{y})$ and the *observed* functional relationship $p(\mathbf{y}|\mathbf{x})$. Notably, this observed mapping may differ from the true system mapping in sparsely explored regions, where it is effectively inferred through extrapolation rather than direct observation.

This setting exemplifies a well-regulated autonomous system as defined in Definition 1: Although the joint distribution and the *observed* mapping evolve over time, the underlying input-output relationship, and thus the mutual information $I(\mathbf{x}; \mathbf{y})$, remains stable.

Expected behavior: Because the system is well regulated, a meaningful surprise measure should not react to exploration-driven distributional shifts. In particular, variations in $p(\mathbf{x})$ and the induced changes in $p(\mathbf{y}|\mathbf{x})$ should not be interpreted as unexpected events. Accordingly, we do not expect MIS to be violated.

As shown in Figure 1, MIS evolves steadily within its expected bounds, reflecting stable learning despite continuous changes in the observed data distribution. In contrast, the single-instance Shannon and Bayesian surprises fluctuate erratically, frequently producing spikes that lack clear interpretability. Their cumulative and rolling variants, although smooth, tend to exhibit largely monotonic trends that offer little actionable insight for decision making. Notably, the cumulative Shannon surprise shows a pivot around 25 sampling steps, which can be attributed to increased coverage of the sampling space: As the model becomes more familiar with the environment, familiar observations begin to dominate and outweigh unseen observations in new samples.

4.1.2. Scenario 2: Overexploitation. In this scenario, the system repeatedly samples a previously seen point from $x \in [0, 30]$, specifically observing the pair $(x, y) = (7, 7)$ 100 times. This simulates stagnation.

Expected behavior: Surprise should diminish as no new information is gained. This mirrors the stagnation case in Section 3.3, and we expect MIS to violate its lower bound.

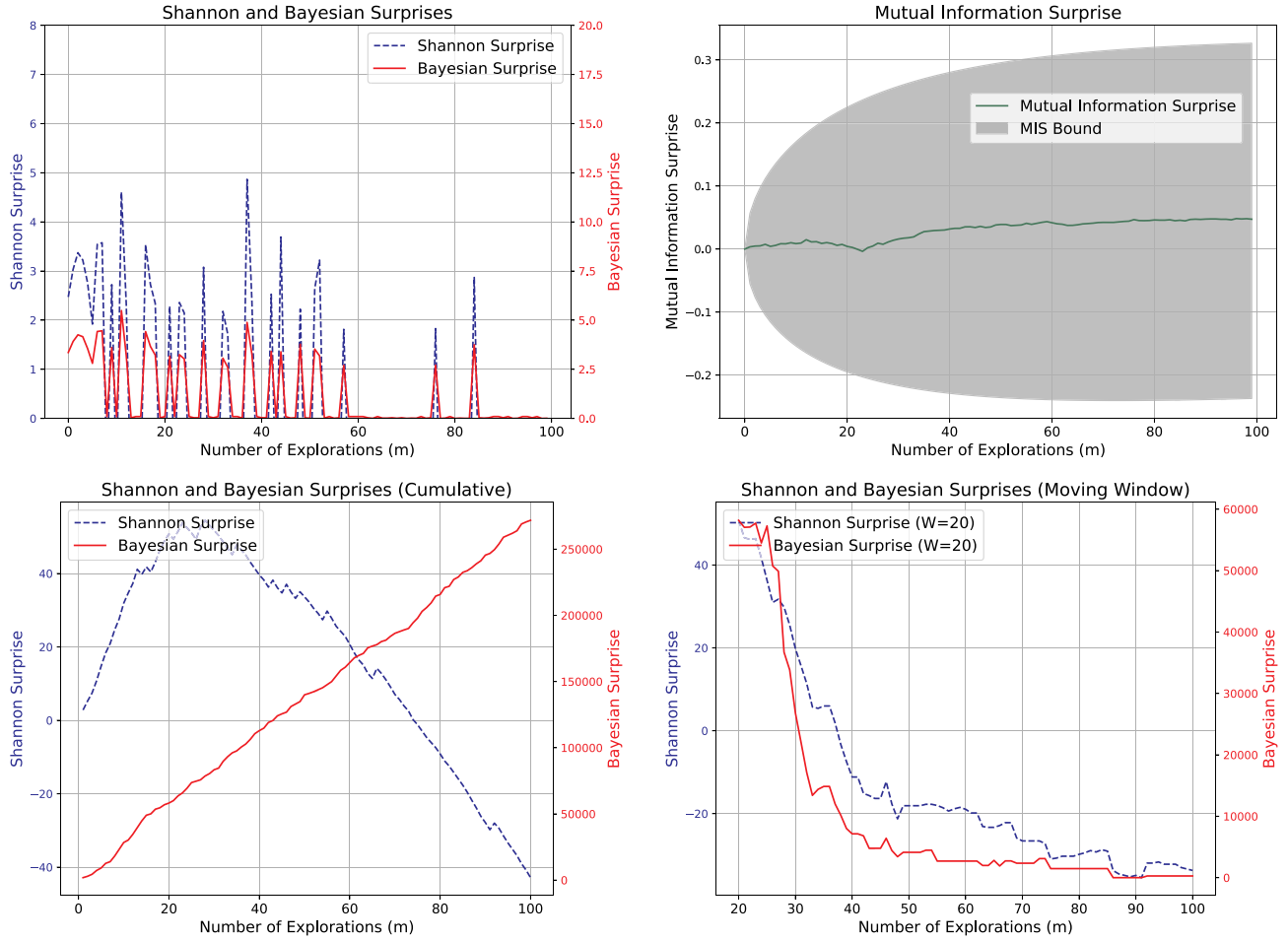
Figure 2 shows that MIS falls below its lower bound, clearly signaling a lack of knowledge gain. Although single-instance Shannon and Bayesian surprises also exhibit a downward trend, they lack a principled lower threshold, limiting their ability to reliably detect such behavior. As noted in Zamiri-Jafarian and Plataniotis (2022) and Ahmed et al. (2025), both Shannon and Bayesian surprises are inherently one-sided measures, which further constrains their interpretability. Similar patterns are observed for their multi-instance variants, with one notable exception: Cumulative Bayesian surprise increasingly labels the exploitation behavior as more surprising over time, whereas rolling Bayesian measures deem it progressively less surprising. Nevertheless, due to its KL divergence formulation, Bayesian surprise lacks a clear probabilistic threshold, rendering this trend difficult to translate into actionable signals.

4.1.3. Scenario 3: Noisy Exploration. We perform standard exploration over $x \in [30, 100]$ but apply random corruption to the outputs $\mathbf{y}$, replacing each with a uniformly random digit between zero and nine. This simulates exploration without informative feedback.

Expected behavior: Despite novel inputs, the system should register confusion if understanding fails to improve. This mirrors the noise-increase case in Section 3.3, and we expect MIS to violate its lower bound.

Figure 3 confirms this pattern: MIS drops below its expected range, accurately signaling knowledge loss. In contrast, Shannon and Bayesian surprises again exhibit erratic behavior without consistent trends. Meanwhile,

Figure 1. (Color online) Surprise Measures During Standard Exploration



the multi-instance Shannon surprise diverges to infinity, and the cumulative and rolling variants of Bayesian surprise display contradictory monotonic patterns, further limiting their interpretability and practical usefulness.

4.1.4. Scenario 4: Aggressive Exploration. This scenario enforces strict exploration over $x \in [30, 500]$, where each new sample is far from all observed points (i.e., outside the ± 1 neighborhood range).

Expected behavior: Aggressive exploration without verification can lead to overconfidence. This mirrors the aggressive exploration case in Section 3.3, and we expect MIS to exceed its upper bound.

Figure 4 shows MIS exceeding its upper bound, consistent with the expected behavior under pure exploration. In contrast, the single-instance Shannon and Bayesian surprises again fluctuate unpredictably. The multi-instance Shannon surprise diverges to infinite values, whereas the cumulative and rolling versions of Bayesian surprise once more exhibit contradictory monotonic trends, limiting their interpretability.

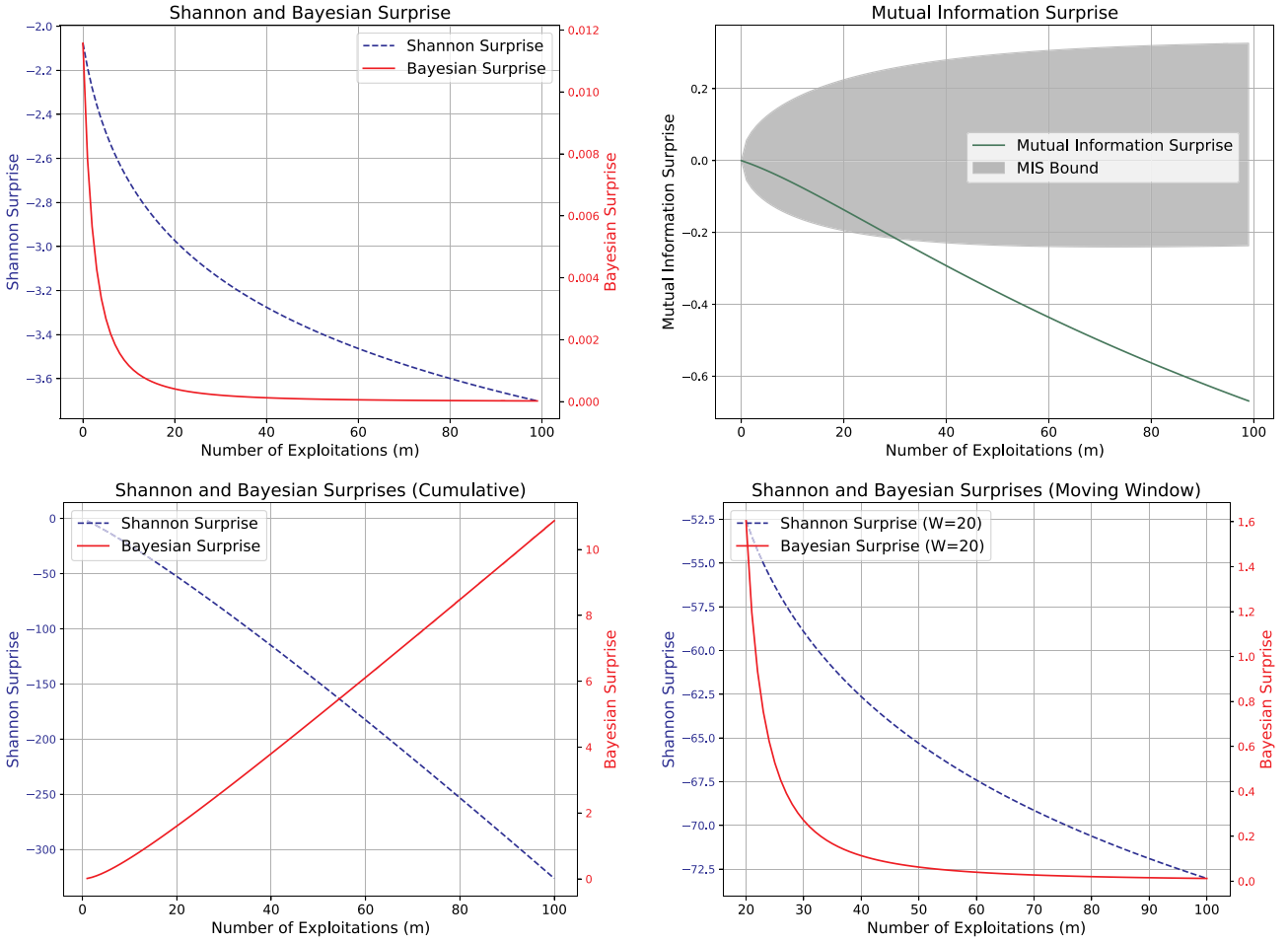
4.1.5. Scenario 5: Noise Decrease. To simulate noise reduction, we begin with 100 initial observations from $x \in [0, 30]$, paired with a randomly assigned output $y \in [0, 9]$. New samples are drawn from the same x range but the new y is produced using the deterministic modulus function in Equation (11).

Expected behavior: Reduced noise implies stronger input-output dependency, and we thus expect MIS to exceed its upper bound.

Figure 5 confirms this: MIS grows beyond its bound. Shannon and Bayesian surprises show similar behaviors to the prior scenarios.

4.1.6. Scenario 6: Discovery of New Output Values. We modify the function in the unexplored region ($x > 30$) to $y = x \mod 10 - 10$, introducing a different behavior while keeping the original function unchanged in [0, 30].

Figure 2. (Color online) Surprise Measures Under Overexploitation



Expected behavior: A competent surprise measure should register this new structure as a meaningful discovery. This mirrors the novel discovery case in Section 3.3, and we expect MIS to exceed its upper bound.

Figure 6 shows MIS sharply exceeding its expected trajectory, signaling successful identification of a structural shift. Shannon and Bayesian surprises again fail to provide consistent or interpretable responses.

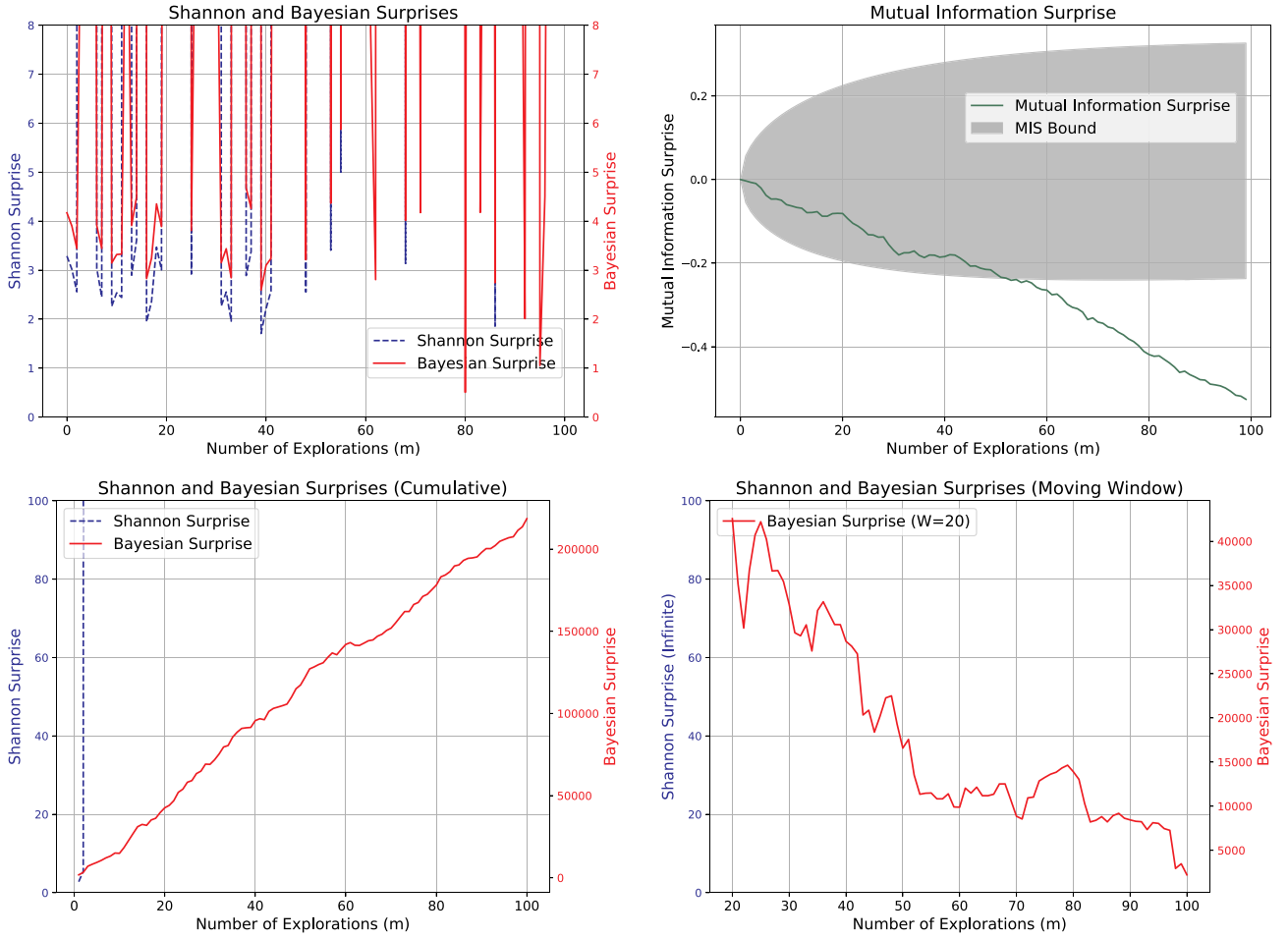
4.1.7. Summary. Across all scenarios, MIS reliably indicates whether the system is genuinely learning, stagnating, or encountering degradation. It responds to the structure and value of observations, more than just novelty. In contrast, single-instance Shannon and Bayesian surprises often react to superficial fluctuations and display numerical instability, and multi-instance Shannon and Bayesian surprises often exhibit simple and even contradictory (cumulative versus rolling) monotone behaviors. Furthermore, the MIS progression bound remains consistent and interpretable across all scenarios, whereas Shannon and Bayesian surprises lack a universal scale or threshold, as reflected by their inconsistent magnitudes across Figures 1–6. This inconsistency limits their effectiveness as a reliable trigger. Overall, this simulation study demonstrates MIS not only as a novel metric for quantifying surprise, but also as a more trustworthy indicator of learning dynamics—making it a promising tool for autonomous system monitoring.

Table 2 summarizes the differences between MIS and the Shannon and Bayesian family of surprises.

4.2. Pollution Estimation: A Case Study

To demonstrate the practical utility of our proposed MIS reaction policy, we apply it to a real-time pollution map estimation scenario. We evaluate the impact of integrating the MIS reaction policy on system performance in a dynamic, nonstationary environment. Specifically, we compare two approaches: a selection of baseline sampling strategies and the same strategies *governed* by our MIS reaction policy.

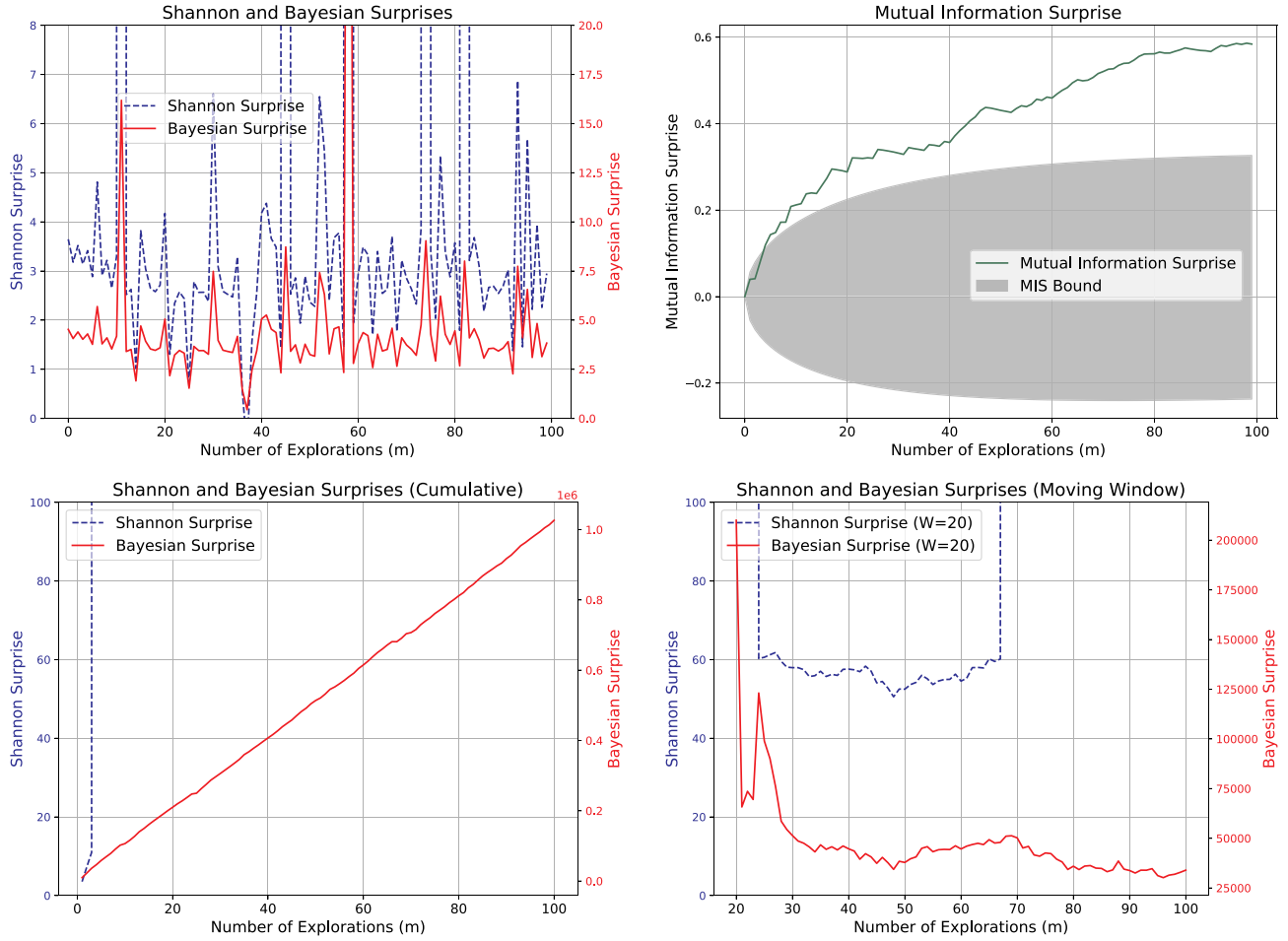
Figure 3. (Color online) Surprise Measures Under Noisy Exploration



4.2.1. Data Set: Dynamic Pollution Maps. We utilize a synthetic pollution simulation data set comprising 450 time frames, each representing a 50×50 pollution grid. Initially, the environment contains three pollution sources, each emitting high pollution at a fixed level. The rest of the field exhibits moderate and random pollution values. Over time, the pollution levels across the entire field evolve due to natural diffusion, decay, and wind effects. Moreover, every 50 frames, a new pollution source is added to the field at a random location. These new sources elevate the overall pollution levels and alter the input-output relationship between the spatial coordinates and the pollution intensity. Figure 7 displays a snippet of the pollution map at two intermediate time points. The simulation details for the dynamic pollution map generation are provided in the Online Appendix.

4.2.2. Sampling Strategies. As discussed in Section 3.4, the MISRP is designed as a data-stream management mechanism that complements existing exploration-exploitation strategies. To evaluate its effectiveness, we conduct simulations using three well-established sampling strategies: the surprise-reactive (SR) sampling method proposed by Ahmed et al. (2025), implemented with either Shannon or Bayesian surprises; the subtractive clustering/entropy (SC/E) active learning strategy introduced by Cebron and Berthold (2009); and the greedy search/query-by-committee (GS/QBC) active learning strategy used in Islam et al. (2025). In addition, input-output data stream management has long been studied in the concept drift and anomaly detection literature. Accordingly, we compare MISRP against one of the most widely used cumulative drift detection metrics, the cumulative Shannon surprise (CSS) defined in Equation (12). Following standard practices in concept drift detection, once CSS exceeds a predefined threshold (CSS > 2.3), corresponding to the product likelihood falling below 0.1, previously collected data are discarded. Note that we do not benchmark the cumulative Bayesian surprise because, unlike the Shannon surprise, the KL divergence nature of the Bayesian surprise makes its triggering threshold ill defined, prohibiting the design of a reaction mechanism.

Figure 4. (Color online) Surprise Measures During Aggressive Exploration



1. **SR:** The surprise-reactive sampling method (Ahmed et al. 2025) switches between exploration and exploitation modes based on the observed Shannon or Bayesian surprise. By default, SR operates in an exploration mode guided by the widely used space-filling principle (Joseph 2016), selecting new sampling locations via the min-max objective:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmax}} \min_{\mathbf{x}_i \in \mathbf{X}} \|\mathbf{x} - \mathbf{x}_i\|_2,$$

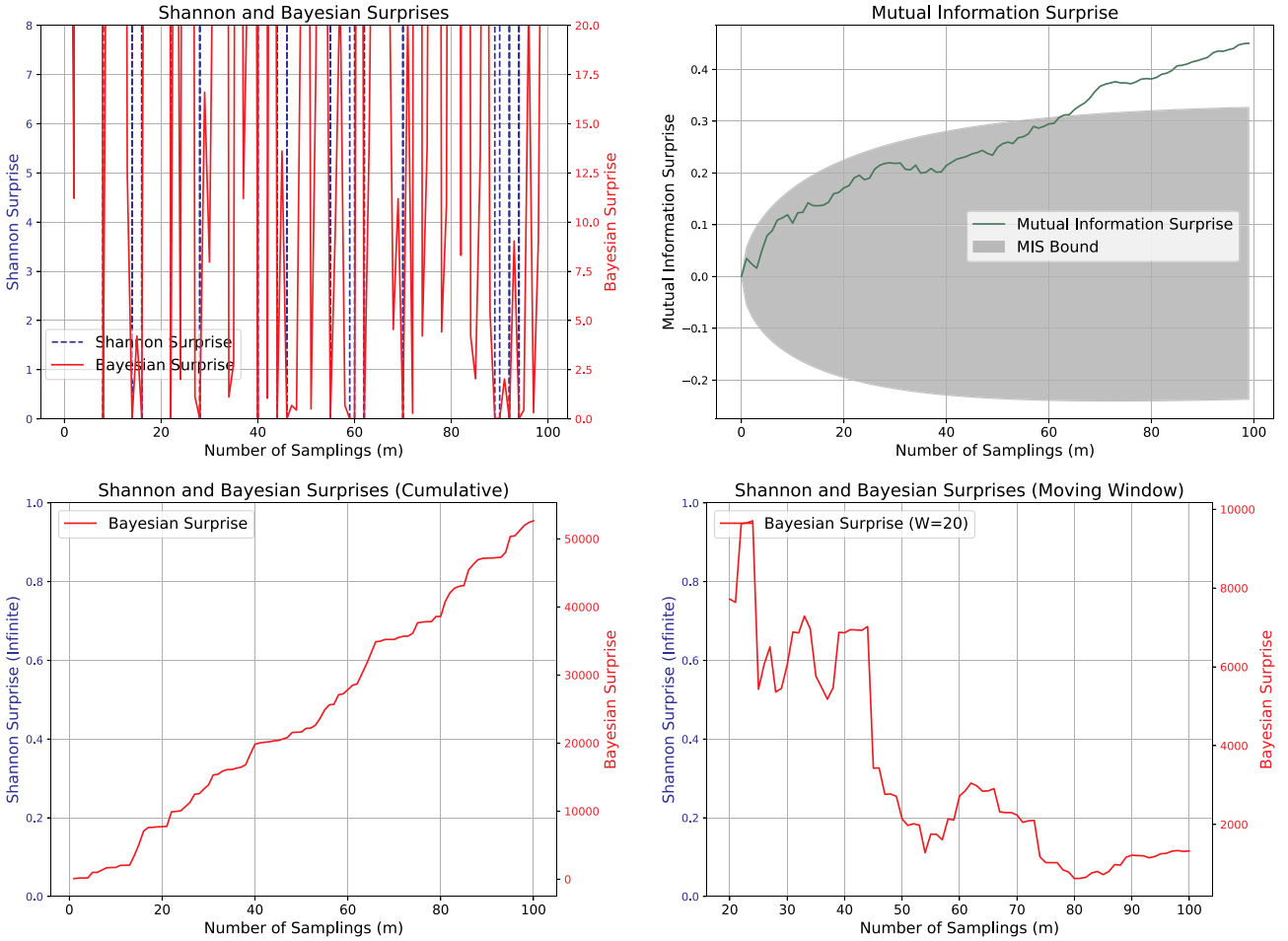
where $\mathbf{X}$ denotes the set of existing observations. Upon encountering a surprising event (in terms of either the Shannon or Bayesian surprise), SR switches to exploitation mode, performing localized verification sampling within the neighborhood of the surprise-triggering location. This continues either for a fixed number of steps defined by an exploitation limit t, or until an unsurprising event occurs. If exploitation confirms that the surprise is consistent (i.e., persistent surprise until reaching the exploitation threshold), all corresponding observations are accepted and incorporated into the pollution map estimation. Conversely, if an unsurprising event arises before the threshold is reached, the surprising observations are deemed anomalous and discarded. For the Shannon surprise, we set the triggering threshold at 1.3, corresponding to a likelihood of 5%. For the Bayesian surprise, we use the postdictive surprise and adopt the threshold of 0.5, following Ahmed et al. (2025).

MISRP: The MISRP modifies SR by dynamically adjusting the exploitation limit t. When increased exploitation is needed, t is incremented by one. For increased exploration, t is decremented by one, with a lower bound of $t = 1$.

2. **SC/E:** The subtractive clustering/entropy active learning strategy (Cebron and Berthold 2009) selects the next sampling location by maximizing a custom acquisition function. For an unseen region $\mathcal{X}$ and a probabilistic predictive function $\hat{f}(\mathbf{x})$ trained on the observed data, the acquisition function is defined as

$$a(\mathbf{x}) = (1-\eta)\mathbb{E}_{\mathbf{x}' \in \mathcal{X}}[e^{-\|\mathbf{x}-\mathbf{x}'\|_2}] + \eta H(\hat{f}(\mathbf{x})),$$

Figure 5. (Color online) Surprise Measures During Noise Decrease



where η is the exploitation parameter, with a default value of 0.5, and $H(\hat{f}(\mathbf{x}))$ denotes the entropy of the predictive distribution at $\mathbf{x}$. A larger value of η emphasizes sampling at locations with high predictive uncertainty near previously seen points, promoting exploitation. A smaller value favors sampling at representative locations in the unseen region, promoting exploration (Cebron and Berthold 2009).

MISRP: The MISRP modifies SC/E by adjusting the exploitation parameter η. For increased exploitation, η is increased by 0.1, up to a maximum of 1. For increased exploration, η is decreased by 0.1, with a minimum of 0.

3. **GS/QBC:** The greedy search/query by committee active learning strategy (Islam et al. 2025) uses a different acquisition function. Given the set of seen observations $\{\mathbf{X}, \mathbf{Y}\}$ and a model committee $\mathcal{F}$ composed of multiple predictive models trained on these data, the acquisition function is defined as

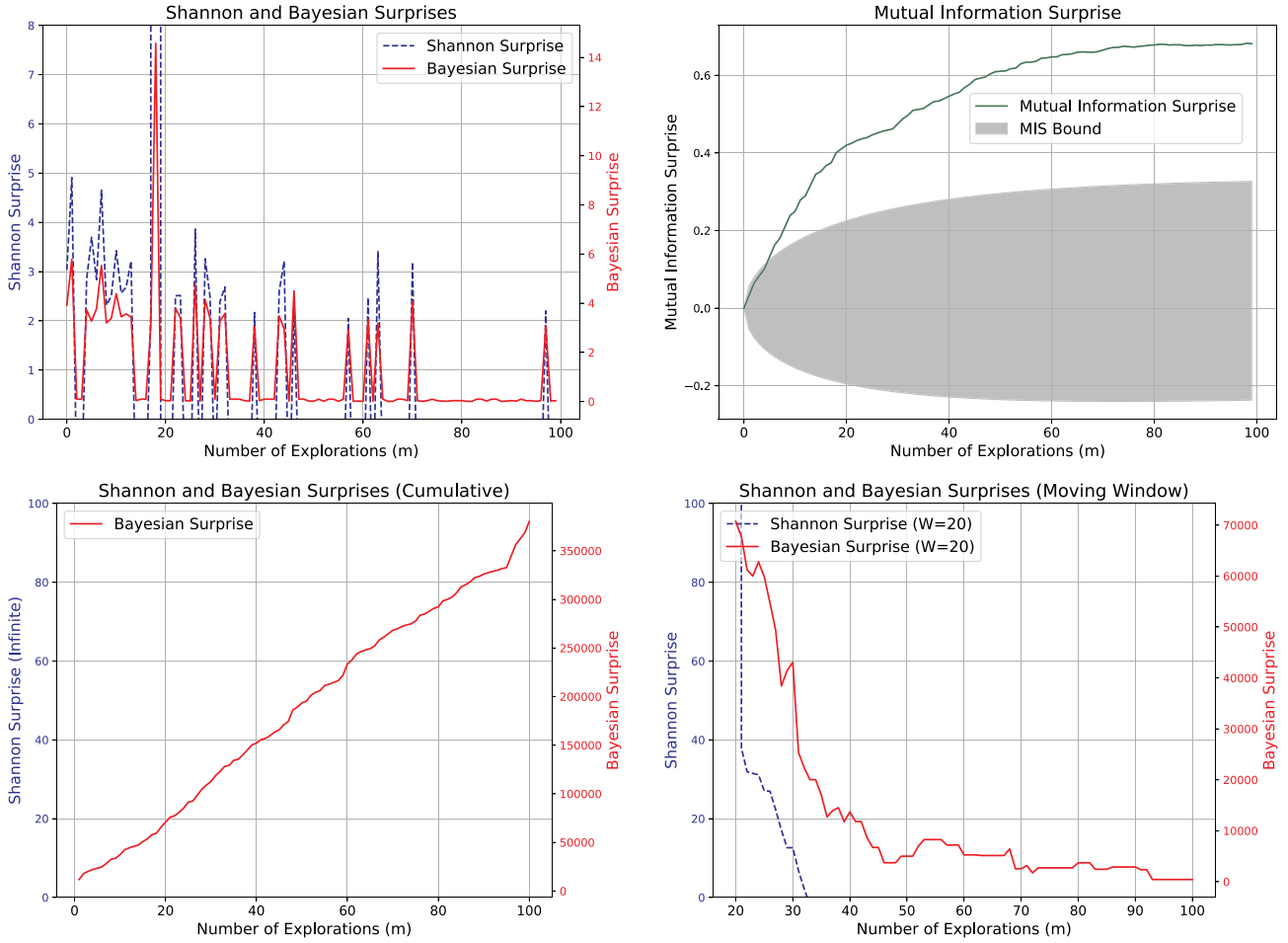
$$a(\mathbf{x}) = (1-\eta) \min_{\mathbf{x}', \mathbf{y}' \in \mathbf{X}, \mathbf{y}} \|\mathbf{x} - \mathbf{x}'\|_2 \|\hat{f}(\mathbf{x}) - \mathbf{y}'\|_2 + \eta \max_{\hat{f}(\cdot), \hat{f}'(\cdot) \in \mathcal{F}} \|\hat{f}(\mathbf{x}) - \hat{f}'(\mathbf{x})\|_2, \tag{13}$$

where the first term encourages exploration by selecting points that are distant from existing observations in both input and output space. The second term promotes exploitation by targeting locations with high disagreement among models in the committee.

MISRP: The MISRP regulates the balance between exploration and exploitation in GS/QBC in the same manner as in SC/E, by adjusting the parameter η.

4.2.3. Experimental Setup. The estimation process is initialized with 10 observed locations uniformly sampled across the pollution field. Each time frame collects 10 new samples according to the chosen sampling strategy, representing the operation of 10 mobile pollution sensors. The pollution field is estimated using a Gaussian process regressor with a Matérn kernel ($\nu = 2.5$) and a noise prior of 10^{-2}, consistently applied across all strategies.

Figure 6. (Color online) Surprise Measures When Exploring a New Region with Novel Outputs



The model predicts pollution levels at specified spatial locations and is updated using both current and historical data, with a maximum of 200 observations retained to reduce computational cost.

For the GS/QBC strategy, the model committee additionally includes regressors with a Matérn $\nu = 1.5$ kernel and a Gaussian kernel with bandwidth 0.1, both using a noise prior of 10^{-2}. These two additional models are used solely for calculating disagreement in Equation (13) and are not employed in pollution map estimation.

The Shannon and Bayesian surprises are computed following the procedure described in Section 4.1. For MIS calculations, we discretize the range of pollution values observed in the data into 100 bins to estimate entropy, and we set the triggering probability at $\rho = 0.1$.

In process forking scenarios, two separate pollution map estimates, $\hat{f}_m$ and $\hat{f}_n$, are produced for subprocesses $\mathcal{P}_m$ and $\mathcal{P}_n$, respectively. The final pollution map estimate is formed as a weighted combination:

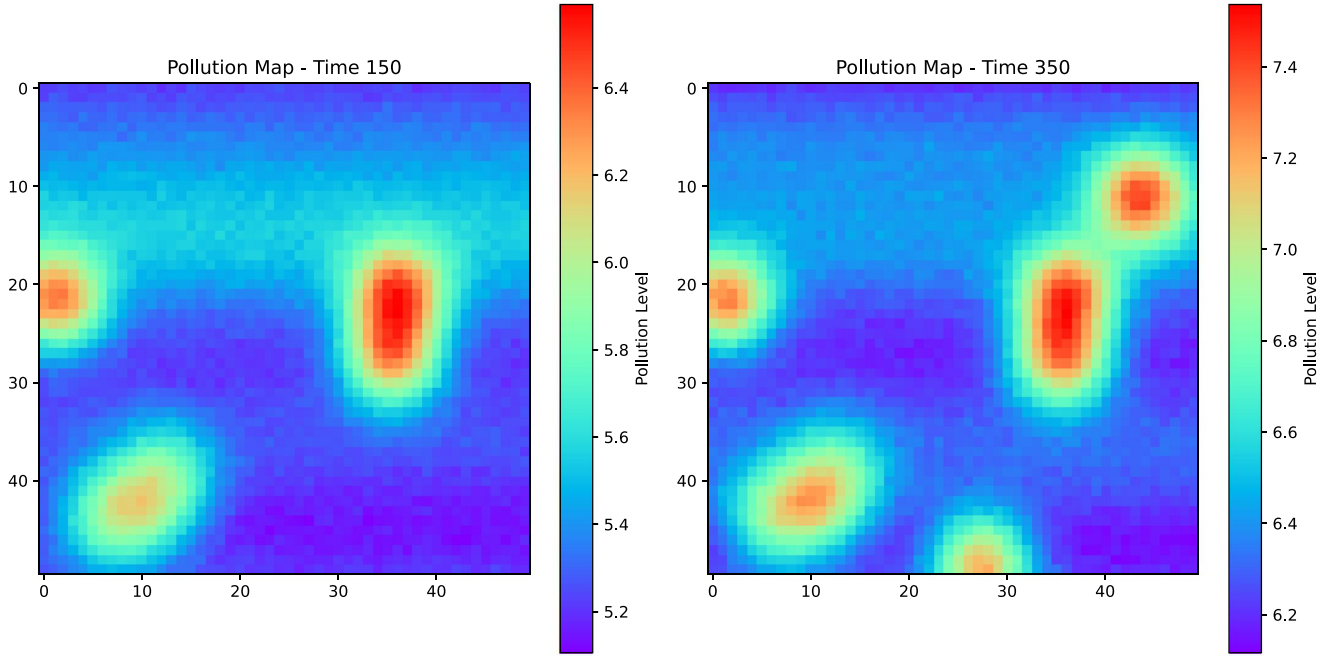
$$\hat{f} = \frac{\sqrt{m}}{\sqrt{m}+\sqrt{n}}\hat{f}_m + \frac{\sqrt{n}}{\sqrt{m}+\sqrt{n}}\hat{f}_n,$$

accounting for generalization errors that scale as $\mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$ and $\mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$, respectively (Chai 2009).

Table 2. Perspective Differences Among Shannon Family Surprises, Bayesian Family Surprises, and MIS

Surprise	Single instance focused	Capture transient changes	Aware of learning progression	Parametric predictive modeling
Shannon Family	✓	✗	✗	✗
Bayesian Family	✓	✗	✗	✓
MIS	✗	✓	✓	✗

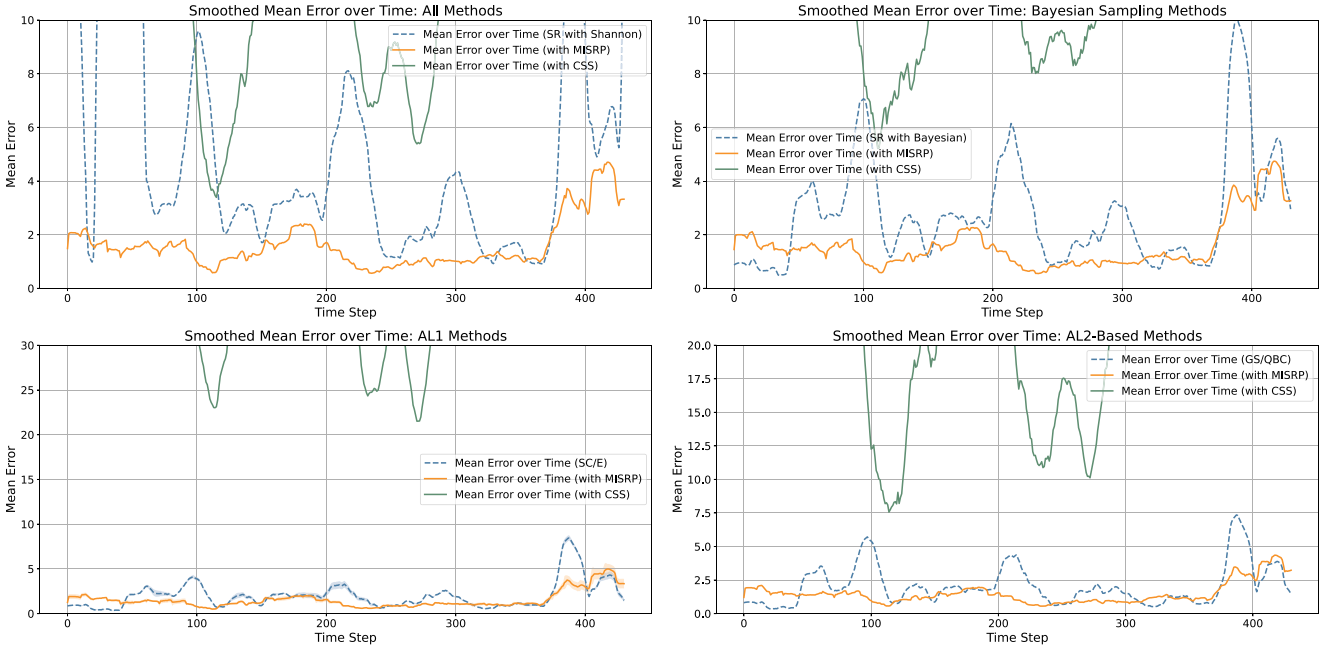
Figure 7. (Color online) Pollution Maps at Time 150 and Time 350



4.2.4. Simulation Results. We assess performance using the mean squared error (MSE) between predicted and true pollution maps at each time step. Because of the dynamic nature of the pollution field, estimation errors exhibit substantial fluctuation. To smooth these variations, we compute a 20-frame moving average of the MSE for both vanilla and MISRP-governed strategies. The results are shown in Figure 8.

Across all comparisons, the baseline strategies display considerable volatility. In contrast, MISRP-governed counterparts produce smoother and consistently lower error curves, highlighting the stabilizing effect of MIS through its ability to facilitate adaptive responses in dynamic environments. CSS on the other hand, is constantly

Figure 8. (Color online) Moving Average Estimation Error over Time



Notes. (Top left) SR with Shannon surprise. (Top right) SR with Bayesian surprise. (Bottom left) SC/E. (Bottom right) GS/QBC.

Table 3. Comparison of Pollution Map Estimation Errors: Baseline vs. MISRP vs. CSS

Sampling strategy	Estimation error (baseline)	Estimation error (CSS)	Estimation error (MISRP)	Mean improvement (over baseline)	Standard error improvement (over baseline)
SR with Shannon	6.64±0.436	19.98±0.645	**1.60±0.043**	76%	90%
SR with Bayesian	2.79±0.096	14.78±0.374	**0.87±0.016**	69%	83%
SC/E	2.02±0.071	75.63±2.371	**1.53±0.045**	24%	36%
GS/QBC	2.07±0.071	35.67±1.205	**1.49±0.039**	28%	45%

Note. Bold entries indicate the best performance.

triggered, leaving very few observations for estimation, thus resulting in much higher estimation error and even higher volatility than the baseline strategies.

Table 3 presents the average estimation errors and their corresponding standard errors. The standard error is measured across 10 Monte Carlo simulations and 450 frames. Across all sampling strategies, incorporating the MIS reaction policy yields a substantial reduction in both mean estimation error and variability. Improvements in estimation error (compared with baseline strategies) range from 24% to 76%, whereas reductions in standard error range from 36% to 90%. On the other hand, CSS shows considerable performance degradation even compared with baseline strategies, highlighting the potential negative impact of traditional concept drift detection policy in this inherently dynamic process. The quantitative results in Table 3 highlight the impact of MISRP on learning performance, demonstrating substantial improvements in estimation accuracy and stability through data stream control.

To further illustrate the advantage of MISRP, we increase the per-frame sampling budget and the initial number of observed locations of the baseline strategies from 10 to 25, and expand the total memory buffer from 200 to 500, in order to assess whether baseline strategies can match the performance of MISRP-governed approaches. Table 4 compares the estimation error of MISRP-governed strategies (maintaining the original sampling budget of 10) against the enhanced baseline strategies. Even with a 2.5× increase in sampling budget, the baseline strategies remain significantly outperformed by their MISRP-governed counterparts.

Thus far, we demonstrated that governing basic sampling strategies with MISRP can substantially enhance learning performance in dynamic environments. To provide a clearer view of how MISRP operates over time, we conduct an additional simulation examining its actions throughout the process.

In this experiment, we simulate a two-phase pollution map evolution governed by the same PDE used in earlier simulation. During the first phase (time 0–250), three pollution sources emit high levels of pollutants, and the map evolves under diffusion, decay, and wind effects. At time step 250, the emission sources are removed, and the decay factor is reduced to 1/20th of its original value. The system then continues evolving for an additional 50 steps.

When the pollution sources exist and are emitting (the dynamic phase, time 0–250), the underlying process is a nonstationary process in which we expect frequent MIS triggering. When the pollution sources are gone (the stationary phase, time 251–300), the pollutants in the area will eventually diffuse to a stationary existence, during which time MIS is expected to stop being triggered.

Figure 9 shows the estimation error progression with action overlays under surprise-reactive sampling based on Shannon surprise. Recall from Section 3.4 that there are two actions employed in MISRP governance: sampling adjustments and process forking. These two actions are marked as vertical lines and shaded regions in the plot, respectively. For clarity, we present the 20-frame moving average of estimation error, whereas the nonsmooth version is provided in the Online Appendix. Actions are displayed 20 steps in advance, corresponding to their first observable effect on the smoothed error trajectory.

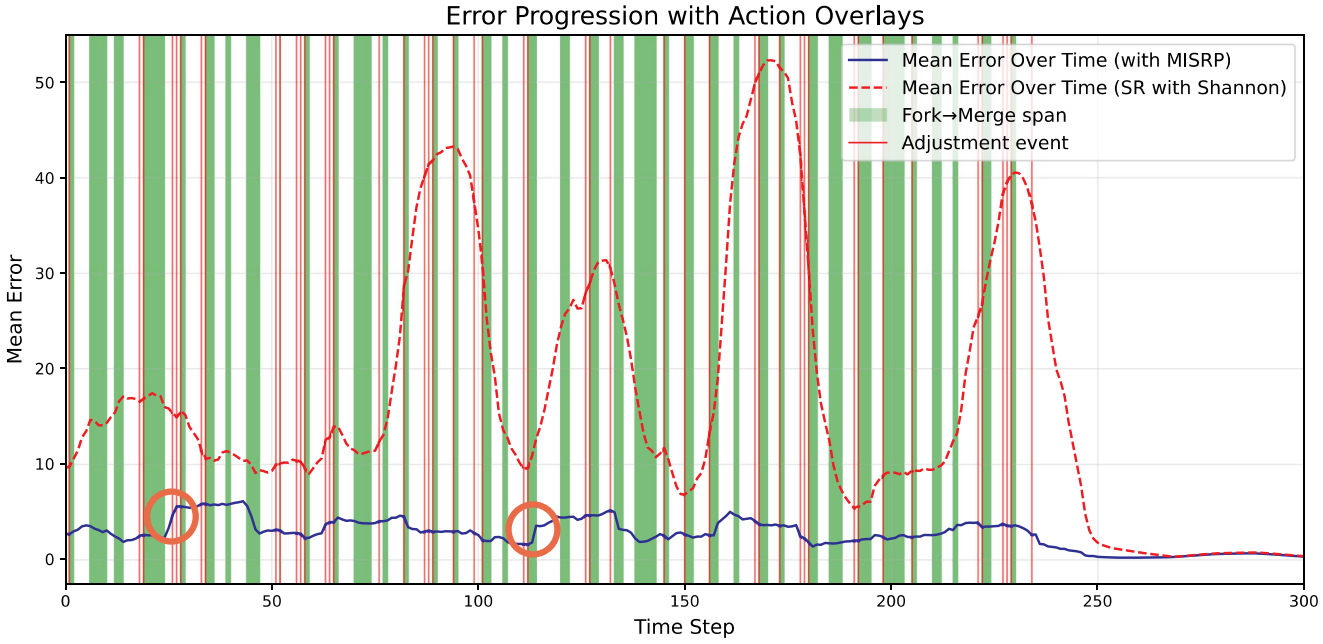
Several key observations emerge from the figure. First, both sampling adjustments and process forking occur frequently during the dynamic phase as expected, highlighting the effectiveness of MISRP's action design in

Table 4. Error Comparison Under Extended Sampling for Baseline Strategies

Sampling strategy	Estimation error (MISRP-governed, budget 10)	Estimation error (Baseline, budget 25)
SR with Shannon	**1.60**	6.23
SR with Bayesian	**0.87**	2.72
SC/E	**1.53**	1.89
GS/QBC	**1.49**	2.00

Note. Bold entries indicate the best performance.

Figure 9. (Color online) Visualization of Estimation Error Progression with MISRP Action Overlays



maintaining low estimation error. Second, sudden spikes in estimation error (circled) under MISRP governance are almost always followed by corrective actions that prevent further error growth, resulting in nonsmooth error progressions after intervention. By contrast, the baseline sampling strategy allows estimation error to rise unchecked. Then, once the system enters the stationary phase, MISRP ceases intervention, aligning with the intuition that a balanced sampling strategy in a *well-regulated* system should not trigger MIS.

5. Conclusion

We started the paper by presenting a vivid picture of the recent race toward autonomous systems. We argued that, although the bodies of various autonomous system have advanced visibly, over the past decade, the brains of these autonomous systems still miss a critical element, which is how to define and react to surprises. Unlike classical surprise measures that characterize statistical irregularities, we reimagine the concept of surprise as a mechanism for fostering understanding. We further define a new surprise metric based on *mutual information* and reframe surprise as a reflection of learning progression grounded in mutual information growth.

We developed a formal test sequence to monitor deviations in the estimated mutual information and introduced a reaction policy, MISRP, that transforms surprise into actionable system behavior. Through a synthetic case study and a pollution map estimation task, we demonstrated that MIS governance offers clear advantages over conventional sampling strategies. Our results show improved stability, better responsiveness to environmental drift, and significant reductions in estimation error. These findings affirm MIS as a robust and adaptive supervisory signal for autonomous systems.

Looking forward, this work opens several promising directions for future research. A natural next step is the development of a *continuous* space formulation of mutual information surprise, enabling its application in large complex systems. Another direction involves designing a *specialized reaction policy*—one that incorporates a sampling strategy tailored directly to the structure and signals of MIS, rather than relying on existing sampling strategies. This could enhance efficiency and responsiveness in highly dynamic or resource-constrained systems. Moreover, pairing MIS with physical probing capability for specific physical systems could unlock the true potential of MIS, as MIS provides new perspectives in system characterization compared with traditional measures.

References

Ahmed I, Bukkapatnam ST, Botcha B, Ding Y (2025) Toward futuristic autonomous experimentation—A surprise-reacting sequential experiment policy. *IEEE Trans. Automation Sci. Engrg.* 22:7912–7926.

Ayed F, Stella L, Januschowski T, Gasthaus J (2020) Anomaly detection at scale: The case for deep distributional time series models. *Proc. Internat. Conf. Service-Oriented Comput.* (Springer, Berlin, Heidelberg), 97–109.

Aytekin C, Ni X, Cricri F, Aksu E (2018) Clustering and unsupervised anomaly detection with l-2 normalized deep auto-encoder representations. *Proc. Internat. Joint Conf. Neural Networks* (IEEE, Piscataway, NJ), 1–6.

Baldi P (2002) A computational theory of surprise. Blaum M, Farrell PG, eds. *Proc. Workshop Honoring Prof. Bob Mceliece on His 60th Birthday: Inform., Coding and Math.* (Springer, Berlin, Heidelberg), 1–25.

Barto A, Mirolli M, Baldassarre G (2013) Novelty or surprise? *Frontiers Psych.* 4:907.

Bayram F, Ahmed BS, Kassler A (2022) From concept drift to model degradation: An overview on performance-aware drift detectors. *Knowledge Based Systems* 245:108632.

Bickel S, Brückner M, Scheffer T (2009) Discriminative learning under covariate shift. *J. Machine Learn. Res.* 10(9):2137–2155.

Bogdoll D, Nitsche M, Zöllner JM (2022) Anomaly detection in autonomous driving: A survey. *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognition* (IEEE, Piscataway, NJ), 4487–4498.

Bondu A, Lemaire V, Boullé M (2010) Exploration vs. exploitation in active learning: A Bayesian approach. *Proc. Internat. Joint Conf. Neural Networks* (IEEE, Piscataway, NJ), 1–7.

Burger B, Maffettone PM, Gusev VV, Aitchison CM, Bai Y, Wang X, Li X, et al. (2020) A mobile robotic chemist. *Nature* 583:237–241.

Çatal O, Leroux S, De Boom C, Verbelen T, Dhoedt B (2020) Anomaly detection for autonomous guided vehicles using Bayesian surprise. *Proc. IEEE/RSJ Internat. Conf. Intelligent Robots Systems* (IEEE, Piscataway, NJ), 8148–8153.

Cebron N, Berthold MR (2009) Active learning for object classification: From exploration to exploitation. *Data Mining Knowledge Discovery* 18:283–299.

Chai K (2009) Generalization errors and learning curves for regression with multi-task Gaussian processes. *Proc. 23rd Adv. Neural Inform. Processing Systems* (ACM, New York), 279–287.

Chang J, Nikolaev P, Carpena-Núñez J, Rao R, Decker K, Islam AE, Kim J, et al. (2020) Efficient closed-loop maximization of carbon nanotube growth rate using Bayesian optimization. *Sci. Rep.* 10:9040.

Cohen K, Zhao Q (2015) Active hypothesis testing for anomaly detection. *IEEE Trans. Inform. Theory* 61(3):1432–1450.

Cover TM (1999) *Elements of Information Theory* (John Wiley & Sons, New York).

Dai T, Vijayakrishnan S, Szczypiński FT, Ayme JF, Simaei E, Fellowes T, Clowes R, et al. (2024) Autonomous mobile robots for exploratory synthetic chemistry. *Nature* 635:890–897.

Dawid AP (1984) Present position and potential developments: Some personal views statistical theory the prequential approach. *J. Roy. Statist. Soc.: Ser. A (General)* 147(2):278–290.

Dinparastdjadid A, Supeene I, Engstrom J (2023) Measuring surprise in the wild. Preprint, submitted May 12, https://arxiv.org/abs/2305.07733.

Doquire G, Verleysen M (2013) Mutual information-based feature selection for multilabel classification. *Neurocomputing* 122:148–155.

Faraji M, Preuschoff K, Gerstner W (2018) Balancing new against old information: The role of puzzlement surprise in learning. *Neural Comput.* 30(1):34–83.

François D, Wertz V, Verleysen M (2006) The permutation test for feature selection by mutual information. *Proc. 14th Eur. Sympos. Artificial Neural Networks* (ESANN, Bruges, Belgium), 239–244.

Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G (2015) Active inference and epistemic value. *Cognitive Neurosci.* 6(4):187–214.

Hou Y, Chen Z, Wu M, Foo CS, Li X, Shubair RM (2020) Mahalanobis distance based adversarial network for anomaly detection. *Proc. IEEE Internat. Conf. Acoustics, Speech and Signal Processing* (IEEE, Piscataway, NJ), 3192–3196.

Islam UJ, Paynabar K, Runger G, Iquebal AS (2025) Dynamic exploration–exploitation trade-off in active learning regression with Bayesian hierarchical modeling. *IISE Trans.* 57(4):393–407.

Itti L, Baldi P (2009) Bayesian surprise attracts human attention. *Vision Res.* 49(10):1295–1306.

Jin S, Deneault JR, Maruyama B, Ding Y (2022) Autonomous experimentation systems and benefit of surprise-based Bayesian optimization. *Proc. Internat. Sympos. Flexible Automation* (ASME, Yokohama, Japan), 1–7.

Joseph VR (2016) Space-filling designs for computer experiments: A review. *Quality Engrg.* 28(1):28–35.

Kamenik JF, Szewc M (2023) Null hypothesis test for anomaly detection. *Phys. Lett. B* 840:137836.

Kolossa A, Kopp B, Fingscheidt T (2015) A computational analysis of the neural bases of Bayesian inference. *Neuroimage* 106:222–237.

Lee MC, Lin JC, Gran EG (2020) Repad: Real-time proactive anomaly detection for time series. *Proc. Internat. Conf. Advanced Inform. Networking Appl.* (Springer), 1291–1302.

Leng J, Zhong Y, Lin Z, Xu K, Mourtzis D, Zhou X, Zheng P, Liu Q, Zhao JL, Shen W (2023) Towards resilience in industry 5.0: A decentralized autonomous manufacturing paradigm. *J. Manufacturing Systems* 71:95–114.

Levinson J, Askeland J, Becker J, Dolson J, Held D, Kammel S, Kolter JZ, et al. (2011) Towards fully autonomous driving: Systems and algorithms. *Proc. IEEE Intelligent Vehicles Sympos* (IEEE, Piscataway, NJ), 163–168.

Liakoni V, Modirshanechi A, Gerstner W, Brea J (2021) Learning in volatile environments with the Bayes factor surprise. *Neural Comput.* 33(2):269–340.

Lian B, Kartal Y, Lewis FL, Mikulski DG, Hudas GR, Wan Y, Davoudi A (2022) Anomaly detection and correction of optimizing autonomous systems with inverse reinforcement learning. *IEEE Trans. Cybernetics* 53(7):4555–4566.

MacLeod BP, Parlane FG, Morrissey TD, Häse F, Roch LM, Dettelbach KE, Moreira R, et al. (2020) Self-driving laboratory for accelerated discovery of thin-film materials. *Sci. Adv.* 6(20):eaaz8867.

Merchant A, Batzner S, Schoenholz SS, Aykol M, Cheon G, Cubuk ED (2023) Scaling deep learning for materials discovery. *Nature* 624:80–85.

Modirshanechi A, Brea J, Gerstner W (2022) A taxonomy of surprise definitions. *J. Math. Psych.* 110:102712.

Montechiesi L, Cocconcelli M, Rubini R (2016) Artificial immune system via Euclidean distance minimization for anomaly detection in bearings. *Mechanical Systems Signal Processing* 76:380–393.

Moreno-Torres JG, Raeder T, Alaiz-Rodríguez R, Chawla NV, Herrera F (2012) A unifying view on dataset shift in classification. *Pattern Recognition* 45(1):521–530.

Nguyen DT, Lou Z, Klar M, Brox T (2019) Anomaly detection with multiple-hypotheses predictions. *Proc. 36th Internat. Conf. Machine Learn.*, Proceedings of Machine Learning Research, vol. 97 (PMLR, New York), 4800–4809.

Nikolaev P, Hooper D, Webber F, Rao R, Decker K, Krein M, Poleski J, et al. (2016) Autonomy in materials research: A case study in carbon nanotube growth. *NPJ Comput. Material* 2:16031.

Paninski L (2003) Estimation of entropy and mutual information. *Neural Comput.* 15(6):1191–1253.

Park HS, Tran NH (2012) An autonomous manufacturing system based on swarm of cognitive agents. *J. Manufacturing Systems* 31(3):337–348.

Prat-Carrabin A, Wilson RC, Cohen JD, Azeredo da Silveira R (2021) Human inference in changing environments with temporal structure. *Psych. Rev.* 128(5):879–912.

Raihan AS, Khosravi H, Bhuiyan TH, Ahmed I (2024) An augmented surprise-guided sequential learning framework for predicting the melt pool geometry. *J. Manufacturing Systems* 75:56–77.

Reis J, Cohen Y, Melão N, Costa J, Jorge D (2021) High-tech defense industries: Developing autonomous intelligent systems. *Appl. Sci.* 11(11):4920.

Rousseeuw PJ, Croux C (1993) Alternatives to the median absolute deviation. *J. Amer. Statist. Assoc.* 88(424):1273–1283.

Schlegl T, Seeböck P, Waldstein SM, Langs G, Schmidt-Erfurth U (2019) F-anoGAN: Fast unsupervised anomaly detection with generative adversarial networks. *Medical Image Anal.* 54:30–44.

Shannon CE (1948) A mathematical theory of communication. *Bell System Tech. J.* 27(3):379–423.

Sugiyama M, Krauledat M, Müller KR (2007) Covariate shift adaptation by importance weighted cross validation. *J. Machine Learn. Res.* 8(5):985–1005.

Szymanski NJ, Rendy B, Fei Y, Kumar RE, He T, Milsted D, McDermott MJ, et al. (2023) An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* 624:86–91.

Verdú S (1994) Generalizing the Fano inequality. *IEEE Trans. Inform. Theory* 40(4):1247–1251.

Wang Y, Miao Q, Ma EW, Tsui KL, Pecht MG (2013) Online anomaly detection for hard disk drives based on Mahalanobis distance. *IEEE Trans. Reliability* 62(1):136–145.

Weller-Fahy DJ, Borghetti BJ, Sodemann AA (2014) A survey of distance and similarity measures used within network intrusion anomaly detection. *IEEE Comm. Surveys Tutorials* 17(1):70–91.

Yurtsever E, Lambert J, Carballo A, Takeda K (2020) A survey of autonomous driving: Common practices and emerging technologies. *IEEE Access* 8:58443–58469.

Zamiri-Jafarian Y, Plataniotis KN (2022) A Bayesian surprise approach in designing cognitive radar for autonomous driving. *Entropy* 24(5):672.

Zhang K, Bui AT, Apley DW (2023) Concept drift monitoring and diagnostics of supervised learning models via score vectors. *Technometrics* 65(2):137–149.

Zhou ZG, Tang P (2016) Continuous anomaly detection in satellite image time series based on z-scores of season-trend model residuals. *Proc. IEEE Internat. Geosci. Remote Sensing Sympos.* (IEEE, Piscataway, NJ), 3410–3413.

Žliobaitė I, Pechenizkiy M, Gama J (2016) An overview of concept drift applications. *Big Data Analysis: New Algorithms New Society* 16: 91–114.