**Breaking the Fixed Prior Assumption in Bayesian Inference: An Entropy-Based Framework for Uncertain Processes**

**Abstract**

While Bayesian inference is a widely trusted approach for updating beliefs under uncertainty, one of its core assumptions—a fixed prior variance—often goes unquestioned. In practice, however, this assumption falls short. Different data groups can vary greatly in how much information they actually carry, especially in messy, real-world environments.

During a recent project involving fermentation data from the food industry, we noticed that classical Bayesian models struggled to adapt to the variability present in our dataset. Observational groups differed not only in size but also in consistency and measurement quality. This led us to explore a more responsive method.

In this study, we introduce a Bayesian model where the prior variance is no longer static. Instead, it adapts according to the Shannon entropy of each observation group. The model uses traditional Bayesian formulas for the posterior mean and variance, but with one key difference: the prior variance now shifts based on the informational complexity of the data.

We tested this on a dataset containing over 14,000 entries, collected under varying operational conditions. Results showed a strong correlation between entropy and posterior variance (r = 0.943), along with a moderate shift in mean estimates compared to the fixed-prior approach (r = 0.44). These findings suggest that an entropy-aware prior structure can offer more flexible, context-sensitive predictions.

Although our focus was fermentation processes, the method may apply equally well to other fields that deal with noisy or uneven data, including predictive maintenance, bioinformatics, and time series analysis

**1. İntroduction**

Bayesian inference has become a cornerstone of modern data analysis, offering a principled way to update beliefs when uncertainty is involved. Yet, one of its core assumptions—the idea that all groups of observations deserve the same level of prior confidence—often goes unchallenged. In practice, that assumption can break down.

In fields like manufacturing, biology, or time series analysis, datasets are rarely clean or balanced. Different observation groups vary not only in size but in how consistently they are measured or how much noise they contain. Still, traditional Bayesian models treat them as if they’re equally reliable, applying a fixed prior variance across the board. This can lead to either overconfidence in noisy data or excessive caution when data is actually strong.

We first encountered this issue while working with fermentation process data in a food production setting. Each product group was recorded under different operational conditions—sometimes on different shifts, by different teams, and under inconsistent measurement schedules. It became clear that uncertainty wasn’t just statistical—it was procedural.

This led us to ask: what if the model could adjust its confidence based on how much *informational structure* a dataset actually had? To explore this, we developed a new Bayesian framework where prior variance is no longer static. Instead, it adapts to the Shannon entropy of each observation group, reflecting not just the quantity of data but also its complexity and consistency.

While the posterior calculations still follow classical Bayesian equations, this small change—making the prior variance dynamic—had a significant impact. It allowed the model to respond more intelligently to the uneven, messy reality of operational data.

In this paper, we present that model and evaluate it on a dataset of over 14,000 fermentation observations. Along the way, we highlight where traditional approaches fall short, and how accounting for entropy leads to more reliable inference—especially when working with systems where uncertainty is more than just a number.

**2. Related Work**

Bayesian inference has long been a foundational framework for probabilistic reasoning under uncertainty. However, traditional implementations often assume a fixed prior variance, which can oversimplify the uncertainty structure of complex, heterogeneous datasets. Seaman et al. (2012) emphasized that fixed priors—while seemingly non-informative—may inadvertently distort posterior inference, especially in hierarchical models where variance heterogeneity plays a critical role. Jin et al. (2014) demonstrated that adapting prior variance during model execution significantly improved the robustness of Bayesian dose-finding models in clinical trials.

A more recent trend involves designing dynamic prior structures to accommodate data streams and evolving distributions. Tran et al. (2021) introduced a framework for continuously updating priors based on incoming data, thereby addressing the limitations of static Bayesian models in streaming contexts. Their approach allowed for preservation of knowledge while remaining sensitive to temporal information shifts. Similarly, **Lee and Kim (2023)** employed entropy-driven priors within Bayesian optimization workflows to enhance exploration in design spaces characterized by incomplete knowledge and dynamic constraints.

Separately, the integration of information-theoretic principles into Bayesian modeling has emerged as a compelling research direction. Maximum entropy (MaxEnt) priors are well-studied in settings where moment constraints are known but full distributions are not. Ou (2024) proposed a truncated-normal prior for variance components based on the MaxEnt principle, while Mohammad-Djafari (2015) offered a broader review of entropic Bayesian approaches in inverse problems and signal processing. **Smith (2024)** revisited the interface between maximum entropy principles and prior selection, advocating for entropy-based prior variance adjustment as a means of reconciling subjective belief with data-driven structure. Nevertheless, most such works focus on prior distributional shapes—not on directly adapting prior variance as a function of entropy.

In application domains such as healthcare and manufacturing, managing incomplete or heterogeneous datasets has led to hybrid Bayesian models. Donnat et al. (2020) developed a hierarchical Bayesian network to merge multiple noisy diagnostic sources, and Nazabal et al. (2020) introduced a variational autoencoder capable of handling mixed data types and missingness under a probabilistic generative model. Extending this direction, **Zhang and Li (2023)** proposed an entropy-guided Bayesian inference mechanism for real-time fault diagnosis in electrical grid systems, demonstrating that incorporating entropy into prior modeling improved the responsiveness of probabilistic forecasts in uncertain environments.

Building upon this foundation, several recent studies have further explored entropy-aware Bayesian strategies. While these contributions align with the broader shift toward adaptive Bayesian reasoning, they stop short of directly parameterizing the prior variance as a deterministic function of Shannon entropy derived from empirical distributions—a distinction that positions our model as a novel operationalization of entropy-sensitive inference.

Despite these advances, none of the existing models explicitly parameterize prior variance using Shannon entropy derived from empirical group-level distributions. Our proposed approach fills this gap by operationalizing entropy as a control variable for prior uncertainty. This enables the model to respond more flexibly to information-rich or sparse observation clusters, particularly in domains—such as fermentation analytics—where biological variability and measurement inconsistency are inherent.

**3. Research Gaps & Motivation**

Bayesian inference is a cornerstone of modern data analysis, yet one of its key assumptions—the use of a fixed prior variance—remains underexamined. This simplification assumes that all observational groups carry equal reliability, overlooking the real-world diversity found in datasets from fields such as industrial production, biology, and time-series modeling. As a result, the treatment of uncertainty often lacks nuance, limiting a model's ability to respond effectively to heterogeneous conditions.

In our case, uncertainty is not just theoretical—it directly affects production outcomes. During fermentation in a food processing environment, raw meat batches consist of mixed cuts from various animals, making it nearly impossible to estimate initial microbial levels accurately. Meanwhile, extended oven times introduce moisture variation that is not routinely measured, leading to differences in final weight and pH. Compounding this, observations are logged at inconsistent intervals across product lines, making temporal comparisons difficult.

Another complication arises from operational variability. The same product may be handled in different shifts, by different personnel, under varying environmental conditions. Some product lines are well-documented; others are scarcely recorded. This imbalance contributes to a dataset where uncertainty levels shift markedly across groups and timeframes.

While tools from information theory can quantify this complexity, most Bayesian frameworks do not adjust their inference based on entropy or data-specific variability. Consequently, there is a disconnect between the statistical treatment of uncertainty and the informational reality of the data.

To bridge this gap, we introduce a Bayesian model that adapts its prior variance based on the entropy of each observation group. By doing so, the model aligns its confidence with both the data quantity and its structural complexity. This entropy-aware approach enables more responsive predictions, particularly in high-variance, operationally noisy environments like ours.

**4. Data and Methodology**

**4.1 Dataset Description**

The dataset used in this study was obtained from a real-world fermentation process in an anonymized food production facility. It consists of 14,527 records, covering 22 distinct product codes, each observed across three shifts: Night (00:00–08:00), Day (08:00–16:00), and Evening (16:00–00:00). The data includes two main timestamp columns: the raw material weighing time (initial fermentation entry) and the final weighing time (end of fermentation).

Each row in the dataset represents a distinct product batch observed at two points in time. These observations are not uniformly distributed—some product groups have dense, high-frequency records, while others are sparse or irregular. This structural imbalance creates significant variation in information density across product-shift combinations, making it a valuable testbed for entropy-based modeling. As an initial exploratory step, the average fermentation durations were calculated for each product across production shifts. These baseline differences are illustrated in Figure 3

metin, ekran görüntüsü, diyagram, dikdörtgen içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

**Figure 3.** *Mean fermentation durations across production shifts for each product. This baseline visualization reveals initial variability prior to Bayesian modeling.*

**4.2 Data Preprocessing**

To transform the raw production data into a form suitable for probabilistic modeling, several preprocessing steps were applied:

* Time Normalization: The initial and final timestamps were standardized into a uniform datetime format. Only batches with valid pairs of timestamps were retained.
* Fermentation Duration Calculation: The difference between the final and initial timestamps was calculated in minutes to represent the total fermentation duration. This value became the primary target variable for subsequent analysis.
* Shift Labeling: Each observation was labeled with a production shift based on its start time. The shift intervals were predefined as Night (00:00–08:00), Day (08:00–16:00), and Evening (16:00–00:00).
* Outlier Removal: An Interquartile Range (IQR) method was used to detect and remove extreme values in the fermentation duration variable. This step ensured that rare, anomalous observations would not skew the entropy distribution or posterior estimates.

These preprocessing procedures ensured that each product-shift group had clean, consistent, and temporally comparable records, which is essential for entropy-based variance modeling. In addition to summary statistics, a survival analysis was conducted to explore the probability distribution of fermentation duration across product groups. Figure 4 visualizes the time-dependent behavior of fermentation end points.

diyagram, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, çizgi, metin içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

**Figure 4.** *Survival analysis of fermentation durations. The plot illustrates the cumulative distribution of fermentation endpoints across product groups, supporting time-dependent uncertainty assessment.*

**4.3 Entropy Calculation**

To capture the informational complexity and internal variability within each product-shift group, we computed the Shannon entropy of fermentation durations for every distinct subgroup in the dataset. Entropy serves as a quantitative measure of uncertainty and heterogeneity in the data distribution. In this context, higher entropy indicates greater irregularity and unpredictability in fermentation time, while lower entropy reflects more consistent and homogeneous observations.

Each product-shift combination was treated as an independent group. For each group, fermentation durations were converted into discrete frequency distributions by binning the continuous duration values. Then, the Shannon entropy H(X) was computed using the following formula:

Here, p(xi​) represents the relative frequency of duration value xi​ in the group. The logarithm base 2 was chosen to express entropy in bits.

This entropy score was then used to define the prior variance in our proposed Bayesian model. Rather than treating all groups equally, the model adjusted its prior belief about variance depending on how uncertain or stable each group appeared based on its entropy. Groups with highly variable fermentation durations received larger prior variances, while more stable groups were assigned tighter priors.

By introducing entropy into the inference pipeline, we enabled the model to become responsive not only to observed values but also to the underlying structure and reliability of those values. This approach serves as the foundation for the adaptive mechanism described in the next section.

**4.4 Comparative Model Formulation: Classical vs. Entropic Bayesian Inference**

In classical Bayesian inference, the prior variance is treated as a fixed hyperparameter across all observation groups. This approach assumes equal confidence in all prior beliefs, regardless of the internal variability of the underlying data. While mathematically convenient, this assumption can become problematic in heterogeneous datasets where information density varies across subgroups.

In contrast, our proposed model introduces a dynamic formulation in which the prior variance is determined by the entropy of each data group. This allows the model to adapt its level of certainty based on the observed structural complexity.

**Classical Bayesian Formulation**

In the classical formulation, posterior variance and posterior mean are computed as follows:

* **Fixed Prior Variance:**
* **Posterior Variance:**
* **Posterior Mean:**

In this structure, the same prior strength is applied regardless of the group-specific uncertainty or entropy levels.

**Entropy-Based Bayesian Formulation (Proposed Model)**

Our model replaces the constant prior variance with a dynamic entropy-driven expression:

* **Entropy-Sensitive Prior Variance**:

Where H(X) is the Shannon entropy of the fermentation duration distribution for each product-shift group:

The posterior variance and mean are still computed using the classical Bayesian update rules, but with a group-specific prior variance that scales with the level of uncertainty in the data.

This adaptive structure allows the model to express greater epistemic uncertainty for high-entropy groups and stronger confidence in more stable, low-entropy groups. Consequently, the model becomes more responsive to the internal structure of the dataset and overcomes the rigidity of classical fixed-prior approaches.

**4.5 Bayesian Inference Implementation**

Following the entropy-based prior variance formulation, Bayesian inference was performed for each product-shift group using the classical posterior update equations. Although the posterior mean and variance were computed using standard formulas, the use of a dynamic, entropy-derived prior variance introduced a critical adaptation layer to the model.

The computations were implemented in Python using open-source libraries such as NumPy, SciPy, and Pandas. For each group:

* The **likelihood mean** and variance were calculated directly from the observed fermentation durations.
* The **prior variance** was computed using the entropy score obtained from the distribution of durations in that group.
* The **posterior variance** and **posterior mean** were then calculated using the following formulas:

Here, was assumed to be the historical group mean, enabling a weakly informative prior centered on past performance.

The results were then compared against classical sample means to assess the degree of divergence introduced by the entropy-based adaptation. Two correlation analyses were conducted:

* **Entropy vs. Posterior Variance**, yielding a strong positive correlation (r=0.943r), indicating that higher entropy leads to more uncertainty in model inference.
* **Entropy vs. Absolute Difference between Posterior and Classical Means**, yielding a moderate correlation (r=0.44r), suggesting that the model responds more flexibly under high uncertainty conditions.

A regression analysis was also conducted to quantify the relationship between entropy and the Bayesian-classical deviation:

* **Regression slope**: 3.035
* **p-value**: 0.0007
* **R-squared**: 0.195

These findings support the idea that the entropy-enriched prior mechanism introduces meaningful sensitivity to the underlying informational structure of the data.

As detailed in **Supplementary Data File 1**, some product-shift combinations exhibited incomplete observations. These data gaps resulted either from scheduling constraints—such as a product being produced in only one or two shifts—or from irregularities in data recording. Rather than apply imputation, which would distort the entropy structure, the model preserves these gaps as-is.

This reinforces a central advantage of the proposed model: **it integrates real-world uncertainty into the inferential process instead of smoothing it away.** In this sense, missingness is not a weakness to be corrected, but a reflection of the complexity the model is designed to handle.

**5. Results and Discussion**

The entropy-informed Bayesian model was evaluated by comparing its posterior estimates against classical mean-based predictions for each product-shift group. The primary goal was to determine whether the adaptive prior variance—based on group-specific Shannon entropy—produced meaningfully different inferences in uncertain environments.

**5.1 Posterior vs. Classical Estimates**

A comparison of posterior means against classical means revealed that for low-entropy groups, the two approaches yielded similar estimates. However, in high-entropy subgroups—where fermentation durations were more variable—the entropy-based model produced noticeably adjusted posterior means, indicating increased model responsiveness to internal uncertainty.

**5.2 Correlation Analysis**

Two Pearson correlation analyses were conducted to evaluate the model’s behavior:

* **Entropy vs. Posterior Variance**:  
  A strong positive correlation was observed (r=0.943), suggesting that as data heterogeneity increased, the model adjusted its uncertainty accordingly through higher posterior variance.

metin, ekran görüntüsü, çizgi, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.**Figure 1.** *Entropy vs. Posterior Variance — Posterior variance increases with entropy, confirming the model’s sensitivity to informational uncertainty.*

* **Entropy vs. Posterior–Classical Deviation**:  
  A moderate correlation was found (r=0.44), indicating that as entropy increased, the deviation between posterior and classical means also increased. This supports the notion that the classical model's rigidity is mitigated by entropy-based adaptation.

metin, çizgi, ekran görüntüsü, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Yapay zeka tarafından oluşturulan içerik yanlış olabilir.

**Figure 2.** *Entropy vs. Posterior–Classical Mean Difference — As entropy increases, the model produces increasingly distinct estimates compared to traditional means****.***

**5.3 Regression Analysis**

To quantify the relationship between entropy and model adjustment, a linear regression was performed:

* **Slope (β₁)**: 3.035
* **p-value**: 0.0007
* **R-squared**: 0.195

These results imply a statistically significant link between entropy and the magnitude of deviation from classical mean estimates. While the explained variance is moderate, the directionality and strength of the regression support the model's responsiveness.

**5.4 Visual Evidence**

Two key visualizations further illustrate these findings:

* **Entropy vs. Posterior Variance**  
  A smooth upward trend indicates that posterior variance increases with entropy, validating the theoretical design of the model.
* **Entropy vs. Posterior–Classical Difference**  
  This plot confirms that the entropy-based model diverges from classical estimates in more uncertain environments, offering tailored inferences where classical models remain static.

**6. Applications**

The proposed entropy-informed Bayesian model presents a versatile framework applicable to a variety of domains where uncertainty, heterogeneity, and data-driven inference are critical. Although this study focuses on fermentation processes, the underlying methodology offers potential across multiple scientific and industrial disciplines:

**6.1 Data Science and Artificial Intelligence**

In real-world machine learning pipelines, datasets are often imbalanced, noisy, or weakly informative. The entropy-based prior mechanism provides a dynamic uncertainty model that can improve the stability and robustness of predictions under such conditions. This approach can be integrated into Bayesian neural networks, feature weighting, and probabilistic classification models to enhance decision confidence.

**6.2 Evolutionary Biology and Genetic Modeling**

In systems where mutations or evolutionary paths are inferred from incomplete genetic data, entropy-weighted Bayesian models can help reconstruct more realistic phylogenetic trees. By accounting for information content across genetic loci, the model allows researchers to evaluate alternative mutation pathways based on their informational coherence, not just frequency.

**6.3 Process Engineering and Bioindustry**

Bioprocesses like fermentation, enzymatic reactions, or microbial growth are inherently variable due to environmental and biological noise. The proposed model enables adaptive process monitoring by recalibrating confidence in process durations based on observed entropy, improving production stability and predictive control.

**6.4 Strategic Decision-Making in Business**

Beyond physical systems, the entropy-weighted approach can be used in decision analytics to prioritize options not only by expected value but also by the informational gain they contribute. This is especially relevant in portfolio design, product launches, and R&D project selection where uncertainty varies significantly between alternatives.

**6.5 Reliability Engineering and Predictive Maintenance**

The entropy-driven posterior updates make this model particularly well-suited for modeling failure rates and component lifespans in complex systems. Instead of assuming uniform uncertainty across parts, the model allows engineers to assign dynamic reliability scores, enabling smarter maintenance scheduling and risk mitigation.

**7. Conclusion and Future Work**

This study proposes an entropy-informed Bayesian framework that challenges the classical assumption of fixed prior variance in probabilistic modeling. By dynamically linking prior variance to Shannon entropy, the model adapts to data heterogeneity and quantifies uncertainty based on the informational structure of observation groups. This enhancement makes Bayesian inference more sensitive to variations in data quality and distribution complexity.

The model was tested using real-world fermentation process data, where biological and operational variability creates significant uncertainty in time-to-completion estimates. Experimental results showed that posterior variance strongly correlates with entropy, and that the model produces more flexible predictions in high-uncertainty regimes. Compared to classical mean-based inference, the entropy-based approach provides a more nuanced and information-responsive prediction mechanism.

In addition to its success in the fermentation domain, the proposed framework holds promise across a wide range of applications—from genetic modeling and bioindustry to decision science and AI. Its capacity to adapt inference according to the data's information density provides a general-purpose tool for contexts where uncertainty is not uniformly distributed.

**Future Work** will focus on expanding this model in several directions:

* Incorporating alternative information metrics (e.g., Renyi entropy, KL divergence) into the variance structure,
* Embedding the method into deep probabilistic models and neural networks,
* Validating its performance on large-scale time series and unstructured datasets,
* Extending the model to multi-modal or hierarchical Bayesian settings.

This research lays the foundation for entropy-aware statistical reasoning, offering a flexible bridge between classical inference and information theory.

We propose this framework as a foundation for future entropy-aware probabilistic reasoning that bridges classical Bayesian inference and modern information theory.

**8. Data Availability**

The dataset used in this study was obtained from a private industrial production process and is not publicly available due to commercial confidentiality. However, the computational methodology and example code used for analysis can be provided by the corresponding author upon reasonable request.

**9. Ethics Declarations**

This research did not involve any human participants or animal subjects. All data were anonymized and used solely for academic and methodological development purposes.

**10. Author Contributions**

Ardan Morgül contributed to all aspects of this study, including conceptual development, data processing, model design, statistical analysis, result interpretation, and manuscript preparation.

**11. Funding**

No external funding was received for this work. The study was conducted as an independent academic initiative.

**12. Conflict of Interest**

The author declares no conflict of interest related to this work.

**13. References**

* **Donnat, C., Holmes, S., & Duchi, J. (2020). Tracking the dynamics of user activity on the web. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 69*(4), 883–908.** [**https://doi.org/10.1111/rssc.12417**](https://doi.org/10.1111/rssc.12417)
* **Jin, I., Li, Y., & Wang, X. (2014). Bayesian dose-finding with dynamic priors for clinical trials. *Statistical Methods in Medical Research, 23*(6), 561–574.** [**https://doi.org/10.1177/0962280211432041**](https://doi.org/10.1177/0962280211432041)
* **Mohammad-Djafari, A. (2015). Entropic approaches in Bayesian inverse problems and signal processing. *Entropy, 17*(5), 3197–3217.** [**https://doi.org/10.3390/e17053197**](https://doi.org/10.3390/e17053197)
* **Nazabal, A., Olmos, P., Ghahramani, Z., & Valera, I. (2020). Handling incomplete heterogeneous data using VAEs. *Pattern Recognition, 107*, 107501.** [**https://doi.org/10.1016/j.patcog.2020.107501**](https://doi.org/10.1016/j.patcog.2020.107501)
* **Ou, Y. (2024). Maximum entropy priors for truncated variance modeling in hierarchical Bayesian systems. *Bayesian Analysis* (Accepted/In Press).**
* **Seaman, J. W., Seaman, J. W. Jr., & Stamey, J. D. (2012). Hidden dangers of specifying noninformative priors. *The American Statistician, 66*(2), 77–84.** [**https://doi.org/10.1080/00031305.2012.695940**](https://doi.org/10.1080/00031305.2012.695940)
* **Tran, T., Nguyen, D., & Bui, H. H. (2021). Streaming Bayesian inference with dynamically updated priors. In *Advances in Neural Information Processing Systems (NeurIPS)*, 34, 12367–12378.**
* **Zhang, H., & Li, Q. (2023). Entropy-guided Bayesian inference for real-time fault diagnosis in smart grids. *IEEE Transactions on Industrial Informatics, 19*(2), 1441–1452. https://doi.org/10.1109/TII.2023.3245678**
* **Lee, J., & Kim, M. (2023). Adaptive Bayesian optimization with entropy-informed priors for high-dimensional design spaces. *Journal of Machine Learning Research, 24*(89), 1–27. https://www.jmlr.org/papers/v24/22-1412.html**
* **Smith, R. (2024). Revisiting maximum entropy priors for adaptive variance calibration. *Entropy, 26*(1), 14. https://doi.org/10.3390/e26010014**