International Journal of Scholarly Research in
Multidisciplinary Studies
Journal homepage: https://srrjournals.com/ijsrms/
ISSN: 2961-3329 (Online)

(Review Article)

Check for updates

Rewriting the tenets: A discourse analysis of probabilistic machine learning in aerospace safety and the turn towards real-time, digital-twin-coupled Bayesian risk inference

Oscar Bediako *

Department of Mathematics and Statistics, College of Engineering and Physical Sciences, University of New Hampshire, New Hampshire, USA.

International Journal of Scholarly Research in Multidisciplinary Studies, 2026, 06(02), 013–018

Publication history: Received on 24 June 2026; revised on 04 August 2026; accepted on 06 August 2026

Article DOI: https://doi.org/10.56781/ijsrms.2026.6.2.0015

Abstract

Probabilistic Machine Learning (PML) is conventionally introduced through a stable set of tenets: quantify epistemic and aleatoric uncertainty, place priors over parameters, marginalise rather than optimise, and let the posterior speak. This paper does not simply apply those tenets to aerospace crash prevention and passenger safety. Instead, it undertakes a discourse analysis of how the *language* of PML is being rewritten as the field migrates from offline model-fitting towards real-time, onboard, digital-twin-coupled Bayesian risk inference. Reading the aerospace-safety literature as a discourse rather than a toolbox, this paper argues that three classical tenets are being quietly inverted: the posterior is no longer a static object of belief but a streaming control signal; the prior is no longer a subjective nuisance but a certifiable engineering artefact; and marginalisation is no longer a purely epistemic act but a temporal, resource-bounded one performed against a hard wall-clock. This shift is formalised with a mixed mathematical core drawing on Bayesian filtering, stochastic hazard modelling and constrained decision theory, and it is shown that the emerging discourse coheres around a single wished-for future: an aircraft that continuously estimates its own probability of catastrophic failure and acts, provably and in bounded time, to reduce it. The contribution is a vocabulary and a set of formal objects for that future, offered as a bridge between the mathematician's instinct for rigour and the safety engineer's instinct for assurance.

Keywords: Probabilistic Machine Learning; Bayesian Filtering; Digital Twin; Aerospace Safety; Hazard Rate; Risk-Bounded Control.

1 Introduction

Every field carries a canonical story about itself. Probabilistic Machine Learning tells a story of humility: where deterministic learning returns a single answer, PML returns a distribution, and where a point estimate conceals its own ignorance, a posterior wears that ignorance openly [1]. This account is customarily taught through four tenets: represent uncertainty explicitly, encode prior knowledge as a distribution, integrate rather than maximise, and update beliefs coherently as data arrive. The distinction between the two irreducible species of uncertainty, aleatoric and epistemic, is now treated as foundational to the discipline [2, 3]. The tenets are correct. They are also, this paper argues, no longer sufficient to describe what practitioners in aerospace safety actually *say* they are doing.

This paper is a discourse analysis. Its object is not a dataset but a way of speaking: the evolving rhetoric through which the aerospace-safety community imports, reshapes and re-legitimises probabilistic learning. Discourse analysis, in the tradition running from Foucault [4] through the sociology of scientific knowledge, asks not only *what* claims are made but *which* claims become sayable, which are foreclosed, and what institutional pressures bend the vocabulary. When a

* Corresponding author: Oscar Bediako.

mathematician reads this literature closely, a striking pattern emerges: the words of PML survive intact while their referents shift. The terms 'posterior', 'prior' and 'marginal' persist as vocabulary, but the safety-critical context loads them with meanings the original tenets never anticipated. The aim here is to make that semantic drift explicit and to give it precise mathematical form.

The choice of aerospace is deliberate. Aviation is the domain in which the cost of a miscalibrated probability is measured in lives, and in which certification authorities demand that every claim of safety be auditable. It is therefore the domain in which the loose, exploratory vocabulary of academic PML collides hardest with an institutional discourse of assurance. The tension is acute because the modern learning tools that make onboard uncertainty estimation feasible, namely variational approximations, Monte-Carlo dropout and deep ensembles, were themselves developed as *approximations*, honest about their own error only in aggregate [5, 6, 7]. Importing an avowedly approximate estimator into a life-critical loop is precisely where the tenets begin to buckle. That buckling is where they are being rewritten.

The aerospace community has, moreover, already articulated the receiving vessel for these tools. The digital twin, a virtual representation of a physical asset that is continuously synchronised with it through data and simulators for real-time prediction, monitoring and control, was named and formalised for aerospace vehicles a little over a decade ago [8] and has since matured into a broad modelling programme [9, 10]. Its defining ambition is to *mirror the life of the flying twin* so as to enable, in the original phrasing of its aerospace proponents, unprecedented levels of safety and reliability [8]. Read as discourse, that ambition is a demand placed on probability theory: to keep a live, individualised estimate of how close *this* airframe, on *this* flight, is to catastrophe. Reliability engineering has long possessed the vocabulary of hazard and survival for populations of components [11]; the digital twin asks that this vocabulary be made conditional and continuous.

As an ambitious methodological programme, this paper does not remain neutral. It champions the future direction that many in the area privately wish for: a real-time, onboard, digital-twin-coupled Bayesian risk estimator that treats the probability of catastrophe as a live quantity to be continuously suppressed. The mathematical scaffolding for such an object already exists in fragments: recursive Bayesian estimation and filtering [12], extreme-value theory for the rare-event tail [13], and chance-constrained control for acting under a probabilistic safety limit [14]. The paper's wager is that reading the discourse closely reveals exactly where these fragments must be joined, and where the missing theorems lie.

1.1 What this paper is not

This is not a survey of PML algorithms, nor a benchmark of neural-network architectures for flight data. It is not a tutorial that re-derives Bayes' theorem for an engineering audience. Such texts exist and apply the standard tenets faithfully [1, 3]. This paper is the deliberate opposite: an argument that the standard tenets, taken at face value, *mislead* about what safety-critical PML has become, and a formal reconstruction of the tenets that the discourse is implicitly adopting instead. A brief note on stance is owed to the reader. The present author writes as a mathematician who finds the wished-for future not merely plausible but mathematically inviting. Discourse analysis has never required the detachment of a pure observer; it requires only that the analyst make the stance explicit so the reader can discount it accordingly.

2 Method: Reading PML as a Discourse

Discourse analysis, in the tradition of Foucault [4], treats a body of text as a system that produces and constrains meaning. Three analytic moves are adapted here to a technical corpus. First, **lexical persistence with referential drift**: terms that remain constant while their operational definitions change are tracked. Second, **foreclosure**: claims that the safety context renders literally unsayable are identified. For instance, the academic sentence 'the posterior is intractable, so calibration may be ignored' cannot be uttered inside a certification dossier. Third, **licensing conditions**: for each legitimate sentence, one asks what mathematical object must exist for it to be honest.

The corpus, treated at the level of rhetorical structure rather than citation, comprises three overlapping literatures: probabilistic inference and Bayesian deep learning [1, 2, 5]; aerospace reliability and prognostics-and-health-management [11]; and digital-twin engineering [8, 9]. These are read not to summarise their findings but to expose the grammar they share. Where a mathematician would normally extract theorems, this paper extracts *speech acts*, and then asks what theorem would have to be true for each speech act to be honest. This is the sense in which the paper's mathematical core is inseparable from its discourse analysis: every rewritten tenet is paired with the formal object that licenses it.

2.1 The First Rewriting: From Static Posterior to Streaming Control Signal

The classical first tenet states that the output of learning is a posterior distribution over unknowns [1]. In the academic register this posterior is an object of contemplation: one reports it, plots its credible intervals and perhaps samples from it. In the aerospace-safety discourse, the posterior has become something no static-belief vocabulary captures: a *signal in a feedback loop*. The sentence that betrays the shift is ubiquitous in the digital-twin literature, where the model is said to track the aircraft in real time [8, 9]. Tracking is not contemplation. Tracking is control.

Formally, let the hidden safety-relevant state of the aircraft at time t be a vector x_t comprising attitudes, structural loads, actuator health and latent degradation modes. The onboard sensors deliver observations y_t. The classical tenet would have us estimate a single posterior $p(x \mid y)$ over a fixed dataset. The rewritten tenet instead demands a *filtering distribution* that evolves with the wall-clock [12]:

$$p(x_t \mid y_{1:t}) \propto p(y_t \mid x_t) \int p(x_t \mid x_{t-1})\, p(x_{t-1} \mid y_{1:t-1})\, dx_{t-1} \,.$$

This is the Chapman–Kolmogorov recursion followed by a Bayes update. Read as discourse, the equation encodes a rewritten tenet: **the posterior is not a conclusion but a state one must maintain.** The integral is a *predict* step (the digital twin propagates belief forward through the flight-dynamics model $p(x_t \mid x_{t-1})$), and the multiplication by the likelihood is a *correct* step against fresh telemetry. The academic posterior has no verb; this one does. It is continuously produced and consumed.

It is worth dwelling on what the recursion forecloses as much as on what it enables. Once the posterior is a maintained state, certain academic sentences become unsayable. One cannot offer 'the model was retrained overnight' as a safety argument, because the object is not trained-then-deployed but perpetually updated in flight. One cannot say 'the posterior is what we believe after seeing the data', because there is no final datum, only the next observation. The very tense of the discourse changes: PML in aerospace speaks in the present continuous. This grammatical shift is not cosmetic. It forces questions the static tenet never raised: does the recursion *forget* old faults at the right rate; is the filter *stable*, so that a transient sensor glitch does not corrupt belief for the remainder of the flight; does the propagated uncertainty neither collapse into false confidence nor explode into uselessness? These are theorems about the filter, and they are demanded by the discourse, not by any single dataset [12].

The consequence for safety is that the object of interest is rarely x_t itself but a functional of the filtering distribution: the instantaneous probability that the state has entered a catastrophic set $C \subset \mathcal{X}$. Define the **live risk**

$$R_t = \mathbb{P}(x_t \in C \mid y_{1:t}) = \int_C p(x_t \mid y_{1:t})\, dx_t \,.$$

In the classical tenet, a probability such as R_t is a summary computed once. In the rewritten tenet, R_t is a *control input*: the aircraft's automation is expected to act so as to hold R_t below a certified threshold ε. The posterior has become a governor. No amount of faithful application of the original tenet reveals this; it appears only when one reads what the engineers are actually asking the distribution to *do*.

2.2 The Second Rewriting: From Subjective Prior to Certifiable Artefact

The second classical tenet is the most philosophically loaded: prior knowledge is encoded as a probability distribution, and its subjectivity is a feature rather than a defect [1]. Academic PML defends the prior against frequentist suspicion by insisting that all inference is conditional on assumptions, so one might as well state them as a distribution. In aerospace, this defence is inadmissible. A certification authority cannot accept 'the prior reflects the analyst's belief' as a safety argument; the conventional certification apparatus is itself rooted in decades of laboratory tests, proto-flight tests and flight histories rather than in stated belief [8]. The discourse therefore performs a remarkable inversion: the prior is stripped of its subjectivity and reconstructed as an *engineering artefact with provenance*.

Concretely, a prior over a degradation-rate parameter θ is no longer justified by belief but by physics-of-failure models, historical fleet data and material-science bounds [11]. The sentence 'a weakly-informative prior was chosen', a badge of honour in academic PML, becomes almost unsayable, because *weakly-informative* reads to a certifier as *unaccountable*. What replaces it is a prior whose every hyperparameter traces to a documented source. The mathematician's task shifts from choosing a convenient conjugate form to constructing a distribution whose support and tails are *defensible*.

This has a precise formal footprint in the tails. Aerospace safety lives in the far tail of the distribution, where academic PML is weakest and least examined. Suppose failure occurs when a load L exceeds capacity, and exceedances above a high threshold u are modelled with a generalised Pareto distribution [13]:

$$\mathbb{P}(L > u + z \mid L > u) = (1 + \xi z / \sigma)^{(-1/\xi)}, \ z \geq 0,$$

with shape ξ and scale σ. The rewritten second tenet insists that ξ, which governs whether the tail is bounded ($\xi < 0$), exponential ($\xi \to 0$) or heavy ($\xi > 0$), cannot be inferred from convenience. A heavy-tailed prior on ξ is not 'conservative by default'; it is a claim about physical reality that must be underwritten by evidence. Here the discourse of assurance disciplines the discourse of Bayesian modelling: the prior becomes a place where physics, not preference, is registered.

The deeper point is that certification reframes the entire prior–likelihood partition. In academic PML the split is a modelling choice; in the safety discourse it is a *division of accountability*. The prior carries what the manufacturer certifies before flight; the likelihood carries what the sensors witness during flight. That the same mathematics supports two utterly different institutional meanings is exactly what discourse analysis is built to reveal.

2.3 The Third Rewriting: From Epistemic Marginalisation to Bounded-Time Integration

The third classical tenet is the mathematician's favourite: integrate over what is unknown rather than optimise it away [1]. Marginalisation is presented as an epistemic virtue, since averaging over parameters yields honest predictive uncertainty. In academic settings the cost of that integral is measured in GPU-hours and tolerated. The aerospace discourse cannot tolerate it, because the integral must complete before the next control cycle. Marginalisation is thus rewritten from an *epistemic* act into a *temporal, resource-bounded* one.

Let τ be the available compute budget per control cycle, perhaps a few milliseconds. The predictive distribution the aircraft actually uses is not the ideal marginal

$$p(y^* \mid y_{1:t}) = \int p(y^* \mid \theta)\, p(\theta \mid y_{1:t})\, d\theta,$$

but an approximation q_τ computable within τ. The very tools that make this feasible onboard, namely Monte-Carlo dropout, weight-uncertainty variational inference and deep ensembles, are explicitly approximations to the intractable posterior [5, 6, 7]. The rewritten tenet says the honest object is therefore not p but the pair (q_τ, δ_τ), where δ_τ bounds the divergence between the approximation and the ideal marginal. A discourse of safety cannot say 'the posterior was approximated' and stop; it must say 'the posterior was approximated, and here is the certified error of that approximation.' This is a genuinely new tenet: **the approximation error of marginalisation is itself a safety-relevant quantity that must be bounded and reported.**

This reframes a large body of technique, from variational inference to sequential Monte-Carlo and ensemble methods, not as competing algorithms but as competing *bargains* between fidelity and time. Where academic PML asks which approximation is most accurate, the safety discourse asks which approximation carries a certifiable δ_τ under a hard deadline. A slower, more accurate method that cannot guarantee completion is, in this discourse, simply unsafe. The mathematician's aesthetic preference for tight bounds is here promoted to an engineering necessity.

3 Synthesis: The Wished-For Object

The three rewritings converge on a single object that the aerospace-safety community, in its more ambitious moments, openly wishes for: an **onboard, digital-twin-coupled Bayesian risk estimator** that runs continuously, holds R_t below threshold, sources its priors from certified physics and reports bounded-time approximation error. This is not a product that exists; it is a horizon the discourse orients towards [8, 9]. Naming it precisely is this paper's central contribution.

The mathematical skeleton of that horizon can be stated compactly. The digital twin maintains the filtering distribution $p(x_t \mid y_{1:t})$ [12]. From it, the live risk R_t is extracted. A decision policy π_t selects an action a_t, such as a trajectory adjustment, an actuator reconfiguration or an alert to the crew, to solve a constrained problem in which the *primary objective is the safety constraint itself* [14]:

$$\text{minimise } \ \mathbb{E}[\,\mathrm{cost}(a_t)\,] \quad \text{subject to} \quad \mathbb{P}(x_{t+1:t+H} \notin C \mid y_{1:t}, a_t) \leq \varepsilon.$$

This is a chance-constrained control problem, and its structure encodes all three rewritten tenets at once. The posterior appears as a streaming input (tenet one). The constraint set C and the threshold ε are certified artefacts, as is the prior baked into the transition model (tenet two). And the probability in the constraint must be evaluated within the control cycle, forcing the bounded-time marginalisation of tenet three. The academic tenets, applied naïvely, would produce none of this; the rewritten tenets produce it necessarily.

It is worth making the hazard-rate view explicit, because it is the language in which reliability engineers already think and thus the bridge along which the wished-for object will actually travel [11]. Let T be the time to a catastrophic event and define the hazard

$$h(t) = \lim_{\Delta \to 0} (1/\Delta)\ \square(t \le T < t + \Delta \mid T \ge t)\ , \quad S(t) = \exp(-\int_0^t h(s)\, ds)\ ,$$

with $S(t)$ the survival function. In the classical reliability discourse, $h(t)$ is a fixed curve estimated from fleet statistics before flight. In the rewritten discourse, the digital twin makes $h(t)$ a *conditional, live* quantity, written $h(t \mid y_{1:t})$, that responds to what this airframe, on this flight, is experiencing now. The survival probability the aircraft is defending is no longer a population average but an individualised, continuously updated estimate. The live risk R_t and the conditional hazard are two faces of the same wished-for capability: to know, at every instant, how close *this* aircraft is to catastrophe, and to act before it arrives.

What separates this horizon from current practice is not a single missing invention but a chain of unclosed guarantees, and the discourse is candid about each link. The filtering distribution onboard today is typically a Gaussian approximation whose validity in the rare-event tail, the only region that matters for safety, is unverified [13]. The priors baked into transition models are often tuned for nominal flight and say little about the off-nominal regimes in which crashes actually originate. The chance constraint is, in current implementations, replaced by conservative surrogates whose conservatism is chosen by hand rather than certified [14]. The discourse names these gaps not as embarrassments but as an agenda: the community *knows* what it wishes it could prove, and it waits for someone to prove it. That waiting is the opening this paper addresses to the ambitious mathematician.

3.1 Why the mathematician is the right author of this future

The three rewritten tenets are not engineering conveniences; they are mathematical demands. A streaming posterior requires a well-posed filtering recursion with provable stability [12]. A certifiable prior requires honest tail modelling and a theory of how prior mass propagates into rare-event probabilities [13]. A bounded-time marginal requires approximation guarantees, that is, divergences with computable, deadline-aware bounds [5]. Each is a place where the discourse *needs* a theorem it does not yet possess. The ambitious mathematician's contribution is not to apply PML's tenets but to supply the missing theorems that make the rewritten tenets honest. That is precisely the future direction the field is wishing into existence.

3.2 Anticipated Objections and Responses

- **'This merely renames Bayesian filtering and chance-constrained control.'** The objects are indeed known; the claim concerns *discourse*, not novelty of equations. The contribution is to show that the vocabulary of PML is being silently reloaded, and to make the new meanings explicit and formal.
- **'Real-time full Bayesian inference is infeasible onboard.'** Correct, and that infeasibility is exactly why the third tenet had to be rewritten. The wished-for object does not require exact inference; it requires approximate inference with a certified error bound δ_τ [5]. The research programme is the pursuit of that bound, not the pretence that the integral is cheap.
- **'Certification will never accept a learned model in the loop.'** This is the deepest tension, and the discourse is aware of it. The rewriting of the second tenet, casting priors as certifiable artefacts and the prior–likelihood split as a division of accountability, is the community's attempt to make learned components *auditable* [8]. Whether authorities accept it is an open institutional question; the mathematics only ensures that the argument, if accepted, is sound.

4 Conclusion

Probabilistic Machine Learning entered aerospace safety wearing familiar clothes: posteriors, priors and marginals. This discourse analysis has argued that beneath the unchanged vocabulary, three tenets have been rewritten. The posterior became a streaming control signal, defended by a filtering recursion. The prior became a certifiable engineering artefact, disciplined by tail geometry and provenance. Marginalisation became a bounded-time act, honest only when paired with a certified approximation error. Together these rewritings point at a single wished-for object: a real-time, onboard, digital-twin-coupled Bayesian risk estimator that holds the live probability of catastrophe below a certified threshold and acts, provably and in bounded time, to keep it there.

For the mathematician, the lesson is liberating rather than deflationary. The field is not asking us to apply settled tenets; it is asking us to author new ones, and to supply the theorems, namely stability of the filter, propagation of prior mass into the tail and deadline-aware approximation bounds, that would make those new tenets trustworthy enough to fly. Reading the discourse closely tells us where the missing rigour lies. Supplying it is how the wished-for future gets built.

References

[1] Ghahramani Z. Probabilistic machine learning and artificial intelligence. Nature. 2015 May 28;521(7553):452-9.

[2] Kendall A, Gal Y. What uncertainties do we need in Bayesian deep learning for computer vision? In: Advances in Neural Information Processing Systems 30 (NIPS 2017). 2017. p. 5580-90.

[3] Hüllermeier E, Waegeman W. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. Mach Learn. 2021 Mar;110(3):457-506.

[4] Foucault M. The archaeology of knowledge. New York: Pantheon Books; 1972.

[5] Gal Y, Ghahramani Z. Dropout as a Bayesian approximation: representing model uncertainty in deep learning. In: Proceedings of the 33rd International Conference on Machine Learning (ICML). PMLR 48. 2016. p. 1050-9.

[6] Blundell C, Cornebise J, Kavukcuoglu K, Wierstra D. Weight uncertainty in neural networks. In: Proceedings of the 32nd International Conference on Machine Learning (ICML). PMLR 37. 2015. p. 1613-22.

[7] Lakshminarayanan B, Pritzel A, Blundell C. Simple and scalable predictive uncertainty estimation using deep ensembles. In: Advances in Neural Information Processing Systems 30 (NIPS 2017). 2017. p. 6402-13.

[8] Glaessgen EH, Stargel DS. The digital twin paradigm for future NASA and U.S. Air Force vehicles. In: 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference; 2012 Apr 23-26; Honolulu, HI. Reston (VA): AIAA; 2012. Paper AIAA 2012-1818.

[9] Rasheed A, San O, Kvamsdal T. Digital twin: values, challenges and enablers from a modeling perspective. IEEE Access. 2020;8:21980-2012.

[10] Tao F, Zhang H, Liu A, Nee AYC. Digital twin in industry: state of the art. IEEE Trans Ind Inform. 2019 Apr;15(4):2405-15.

[11] Rausand M, Høyland A. System reliability theory: models, statistical methods, and applications. 2nd ed. Hoboken (NJ): Wiley; 2004.

[12] Särkkä S. Bayesian filtering and smoothing. Cambridge: Cambridge University Press; 2013.

[13] Coles S. An introduction to statistical modeling of extreme values. London: Springer; 2001.

[14] Blackmore L, Ono M, Williams BC. Chance-constrained optimal path planning with obstacles. IEEE Trans Robot. 2011 Dec;27(6):1080-94.