Offline data assimilation for uncertainty-aware degradation prognostics in systems with sparse data

Edgar Jaber[†]

[†]Université Paris-Saclay, ENS Paris-Saclay, Centre Borelli

ETH Zürich - Structural Mechanics and Monitoring Group Seminar

20/03/2026

Outline

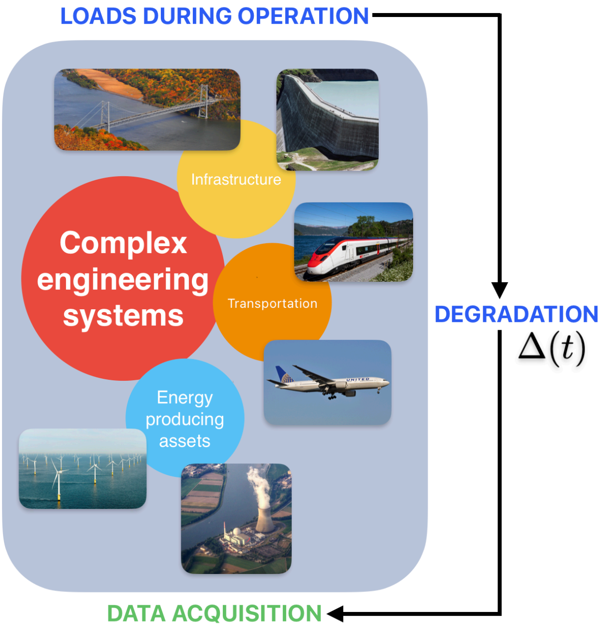

Complex engineering systems
Infrastructure
Transportation
Energy producing assets

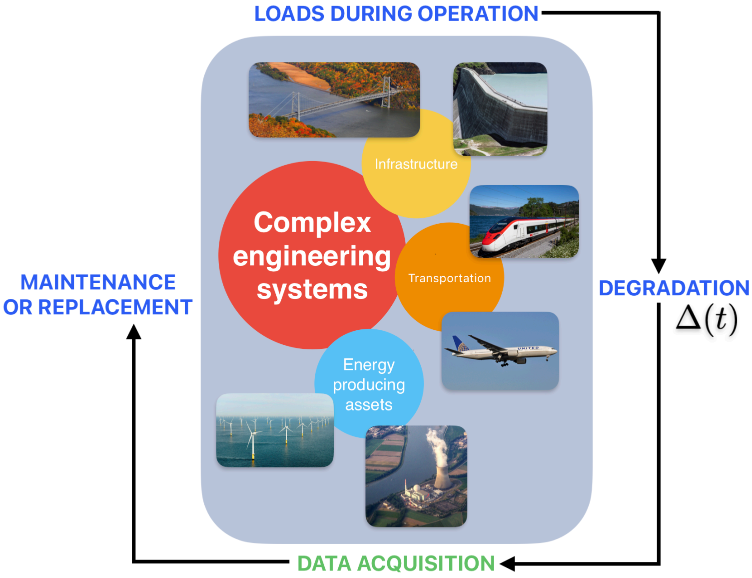
LOADS DURING OPERATION

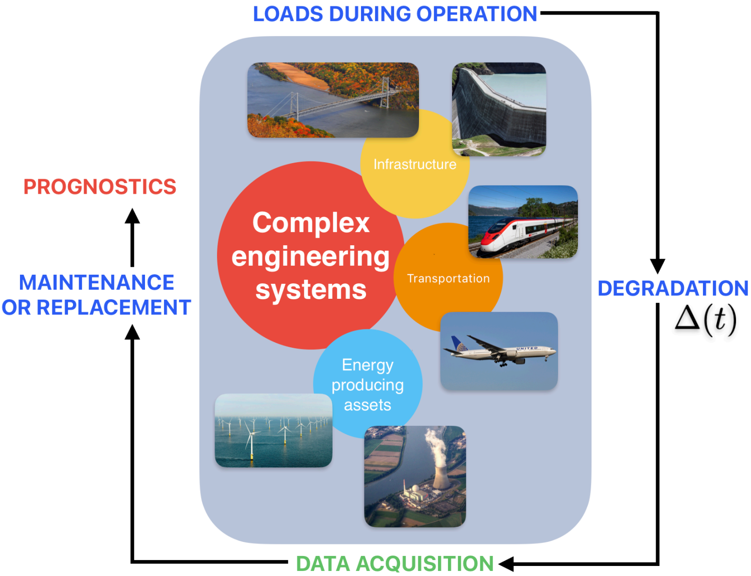
LOADS DURING OPERATION

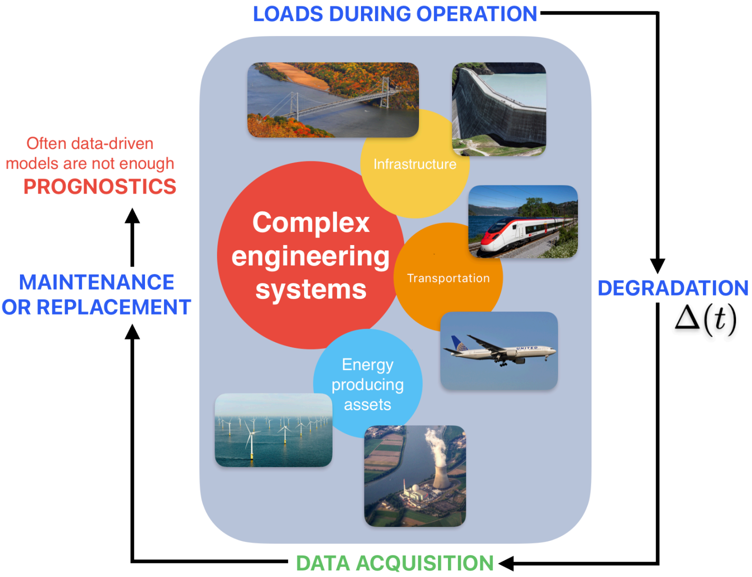
DEGRADATION

$$\Delta(t)$$

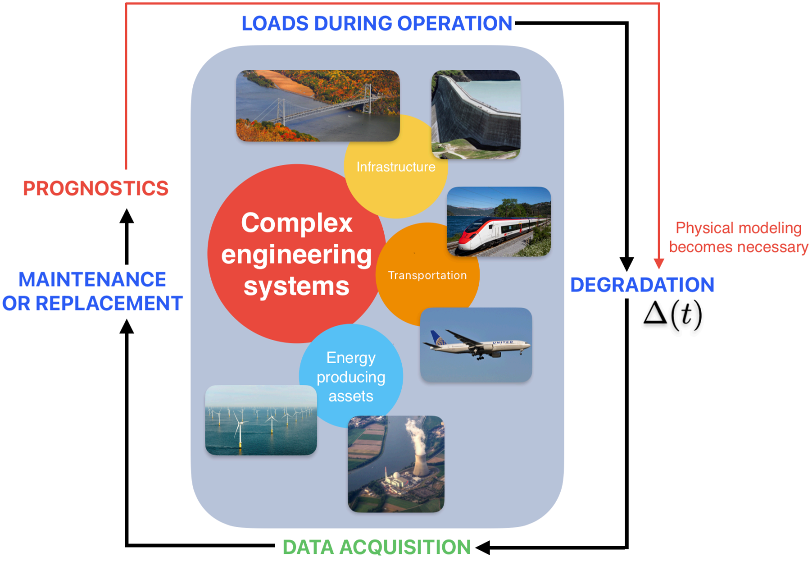

LOADS DURING OPERATION
Complex engineering systems
Infrastructure
Transportation
Energy producing assets
DEGRADATION
$\Delta(t)$
DATA ACQUISITION

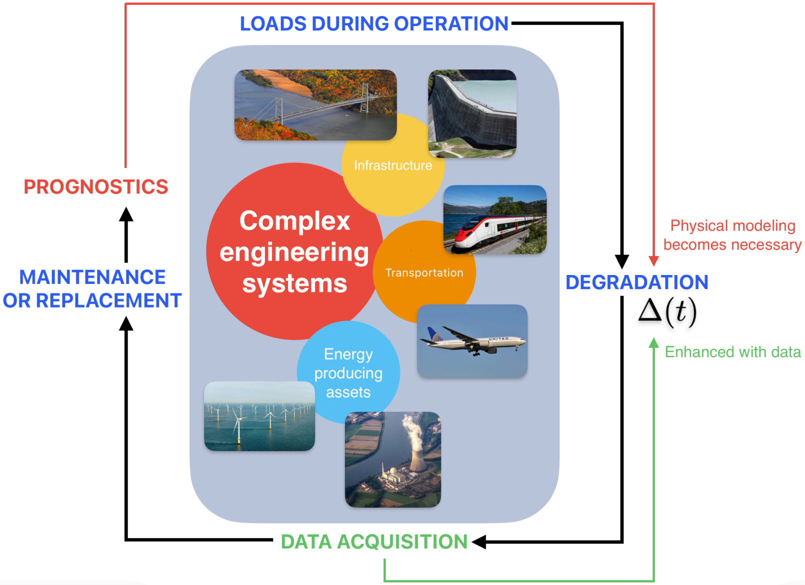

LOADS DURING OPERATION
Infrastructure
Complex engineering systems
Transportation
Energy producing assets
DEGRADATION
$\Delta(t)$
MAINTENANCE OR REPLACEMENT
DATA ACQUISITION


LOADS DURING OPERATION
PROGNOSTICS
MAINTENANCE
OR REPLACEMENT
Complex
engineering
systems
Infrastructure
Transportation
Energy
producing
assets
DEGRADATION
$\Delta(t)$
DATA ACQUISITION

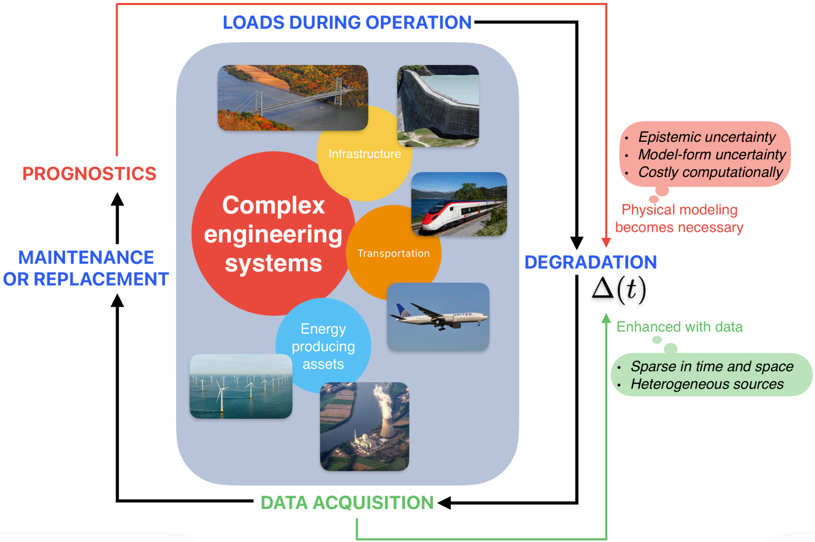


LOADS DURING OPERATION
Often data-driven
models are not enough
PROGNOSTICS
Infrastructure
Complex engineering systems
Transportation
MAINTENANCE
OR REPLACEMENT
DEGRADATION
$\Delta(t)$
Energy producing assets
DATA ACQUISITION


LOADS DURING OPERATION
Infrastructure
Complex engineering systems
Transportation
Energy producing assets
PROGNOSTICS
MAINTENANCE OR REPLACEMENT
Physical modeling becomes necessary
DEGRADATION
$\Delta(t)$
DATA ACQUISITION


LOADS DURING OPERATION
Infrastructure
Complex engineering systems
Transportation
Energy producing assets
PROGNOSTICS
MAINTENANCE OR REPLACEMENT
Physical modeling becomes necessary
DEGRADATION
$\Delta(t)$
Enhanced with data
DATA ACQUISITION

What challenges?



- Degradation level of the system $(t \mapsto \Delta(t))$ → predict the remaining useful life (RUL) [Vachtsevanos et al., 2006] for a fixed threshold $\Delta_* \in \mathbb{R}_+$:

$$\text{RUL}(\Delta_*) = \underset{t > t_P}{\text{argmin}}\{\Delta(t) \geq \Delta_*\} \quad (1)$$

where t_P is the present time.

  → Relies usually on physics-based simulation codes, and/or data driven methods

- Idea: build a hybrid framework for evaluating the SG clogging RUL using the physics-based model and the data-driven models

→ *Data assimilation*, but not straightforward!

Available tools

- *Physics-based computer simulation model* $g : \mathcal{X} \subset \mathbb{R}^d \to \mathbb{R}^N$ with prior uncertainty on input variables $\boldsymbol{X}$, one input $\boldsymbol{x}_0$ gives a full degradation trajectory:

$$g(\boldsymbol{x}_0) = (g(t_1, \boldsymbol{x}_0), \ldots, g(t_N, \boldsymbol{x}_0)) \tag{2}$$

and $\mathrm{pr}_\ell \circ g(\boldsymbol{X}) := g(t_\ell, \boldsymbol{X})$ models $\Delta(t_\ell)$ → grey-box, physics known with non-intrusive surrogate modeling strategy $\widehat{g}$

- q *heterogeneous degradation data groups* (from different sensors, statistical models,...) $\mathcal{D} = (\boldsymbol{y}^1, \ldots, \boldsymbol{y}^q)$ with different sizes $\boldsymbol{y}^i \in \mathbb{R}^{m_i}$ → corresponding to different time indices in $\mathcal{J}_i$ so that $\mathcal{J} = \cup_{i=1}^q \mathcal{J}_i$ and $|\mathcal{J}| = m_1 + \ldots + m_q$, such that:

$$y^i(t_\ell) = \Delta(t_\ell) + \eta_\ell^i \tag{3}$$

with $\eta_\ell^i \sim \mathcal{N}(0, R^i)$ → homoskedastic noise for each data group

Outline

Offline data assimilation

- Perform offline sequential data assimilation over time windows [Geir Evensen, 2022] using ensemble methods
- Let $\boldsymbol{Z} = (\boldsymbol{X}, \boldsymbol{Y})$ where $\boldsymbol{Y} = g(\boldsymbol{X}) + \epsilon = \widehat{g}(\boldsymbol{X})$ are the computer model emulator outputs on a prescribed time window and $\boldsymbol{X}$ are uncertain parameters that do not vary over time
- The goal of assimilation is to estimate the distribution $p(\boldsymbol{Z} \mid \mathcal{D})$
- Use *modular* approach on each time window:
  1. Tailored Bayesian model updating (BMU) to obtain a posterior distribution $p_{\boldsymbol{X}|\mathcal{D}}$ for the input parameters: $(\boldsymbol{X}|\mathcal{D}) \sim p_{\boldsymbol{X}|\mathcal{D}}$.
  2. Apply smoothing to estimate $p(\boldsymbol{Y}|\boldsymbol{X} \sim p_{\boldsymbol{X}|\mathcal{D}}, \mathcal{D})$, yielding the assimilated posterior approximation:

$$p(\boldsymbol{Z}|\mathcal{D}) \approx p(\boldsymbol{Y}|\boldsymbol{X} \sim p_{\boldsymbol{X}|\mathcal{D}}, \mathcal{D}) p_{\boldsymbol{X}|\mathcal{D}}. \tag{4}$$

Illustration of the methodology

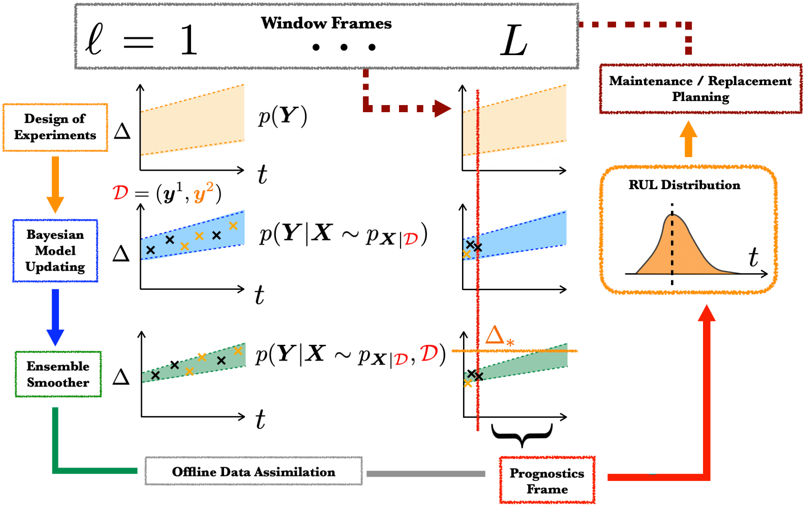


Figure 1: A sketch of the offline data assimilation methodology

On the prognostics window

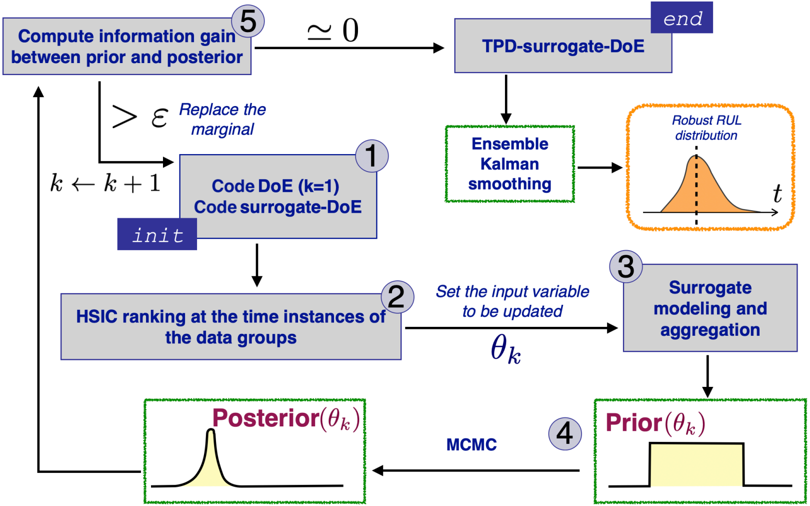
- Assimilation is performed offline until the prognostics window $\ell = L$ is reached, then the RUL distribution can be computed:

$$\mathbb{P}(\mathrm{RUL}(\Delta_*) \leq t_j \mid \mathcal{D}) = \int_{\mathbb{R}} \mathbf{1}\{\mathrm{pr}_{j+1}(\boldsymbol{y}) \geq \Delta_*\} p(\boldsymbol{y}|\mathcal{D}, \boldsymbol{X} \sim p_{\boldsymbol{X}|\mathcal{D}}) d\boldsymbol{y}, \tag{5}$$

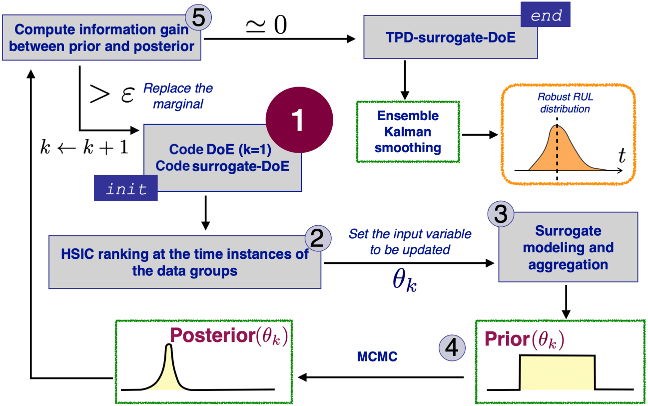
- In practice this probability is estimated using a Monte Carlo ensemble $\{(\boldsymbol{X}^{(i)}, \widehat{g}(\boldsymbol{X}^{(i)}))\}_{i=1}^{n} \sim p_{\boldsymbol{X}|\mathcal{D}} \otimes \widehat{g}\# p_{\boldsymbol{X}|\mathcal{D}}$:

$$\mathbb{P}(\mathrm{RUL}(\Delta_*) \leq t_j \mid \mathcal{D}) \approx \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\{\mathrm{pr}_{j+1} \circ \widehat{g}(\boldsymbol{X}^{(i)}) \geq \Delta_*\} \tag{6}$$

The BMU algorithm



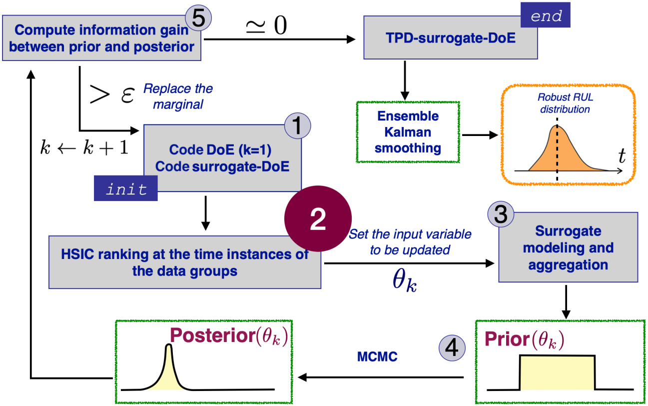
Step (1)



Perform k iterations where $1 \leq k \leq d$:

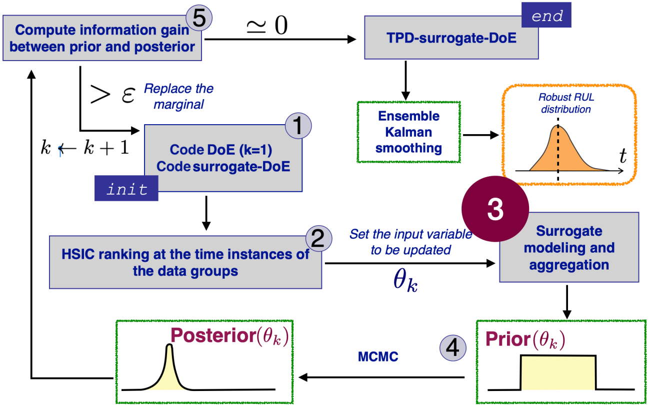
1. If $k = 0$, assume uniform independent priors $\mu_{\boldsymbol{X},k} \simeq \mathcal{U}[-1, 1]^{\otimes d} \to$ generate a design of experiments $\text{DoE}_g^{\mu_{\boldsymbol{X},k}} = \{(\boldsymbol{X}^{(j)}, g(\boldsymbol{X}^{(j)}))\}_{1 \leq j \leq n}$

Step 2



2. Compute **HSIC indices** [Gretton et al., 2005] between input variables and outputs at data time instances → **given data** sensitivity analysis method to assess *individual* input variable influence on the output

Step 3



3. If g is time-costly, build and validate p metamodels $\widehat{\boldsymbol{g}} = (\widehat{g}^{(1)}, \ldots, \widehat{g}^{(p)})$ with chosen strategy $\rightarrow$ avoid metamodeling bias with *convex aggregation* on the unit-simplex chosing $\boldsymbol{w} \in \Delta^{p-1} := \{\boldsymbol{w} \in [0,1]^p, \ \|\boldsymbol{w}\|_1 = 1\}$, fix nominal value of $\boldsymbol{U}_{0,k} = \boldsymbol{u}_{0,k}$ by taking the mean

Metamodeling step 3

The metamodeling process involves:

- *Data generation:* Using the DoE of g at n input samples $\{\boldsymbol{X}^{(i)}\}_{i=1}^{n} \sim \mu_{\boldsymbol{X}}$, assemble the data matrix:

$$\boldsymbol{Y} = \left[g(\boldsymbol{X}^{(1)}), \ldots, g(\boldsymbol{X}^{(n)})\right] \in \mathbb{R}^{N \times n} \tag{7}$$

- *Dimensionality reduction:* Apply a Karhunen–Loève (KL) decomposition [Sullivan, 2015] using the empirical covariance matrix $\widehat{C} = \frac{1}{n}\boldsymbol{Y}\boldsymbol{Y}^{\top}$. Perform singular value decomposition (SVD):

$$\boldsymbol{Y} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{W}^{\top}, \tag{8}$$

  where $\boldsymbol{V}$ contains the KL modes $\{\boldsymbol{\Phi}_k\}_{k=1}^{m}$, and $\boldsymbol{\Sigma}$ holds the singular values.

- *Mode selection:* Retain m modes to capture a prescribed variance (e.g., 99%). Project trajectories onto the retained modes:

$$\xi_k(\boldsymbol{X}^{(i)}) = g(\boldsymbol{X}^{(i)})^{\top}\boldsymbol{\Phi}_k, \quad k = 1, \ldots, m \tag{9}$$

Metamodeling step 3

- *Surrogate modeling:* For each mode k, construct a surrogate model $\widehat{\xi}_k(\boldsymbol{X})$ using a Gaussian process [Rasmussen and Williams, 2006] with identical prior mean and kernel for all modes
- *Reconstruction:* Reconstruct the full trajectory using the surrogate models:

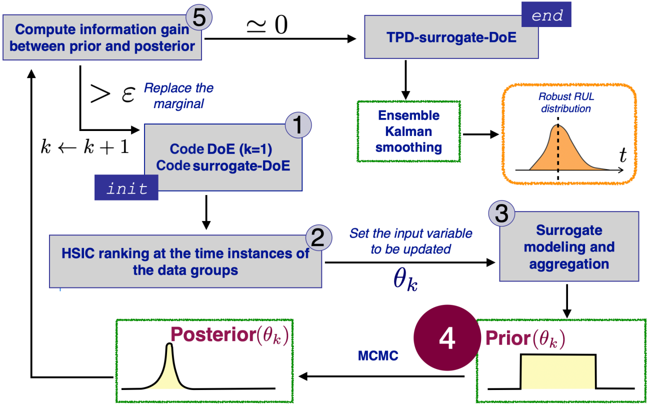
$$\widehat{g}(\boldsymbol{X}) = \sum_{k=1}^{m} \widehat{\xi}_k(\boldsymbol{X}) \boldsymbol{\Phi}_k \tag{10}$$

- *Aggregation:* Combine multiple surrogate models $\{\widehat{g}^{(i)}\}_{i=1}^{p}$ using convex aggregation weights $\boldsymbol{w} \in \boldsymbol{\Delta}^{p-1}$ to form the aggregated surrogate model:

$$\widehat{g}^{\text{agg}}(\boldsymbol{X}) = \sum_{i=1}^{p} w_i \widehat{g}^{(i)}(\boldsymbol{X}) \tag{11}$$

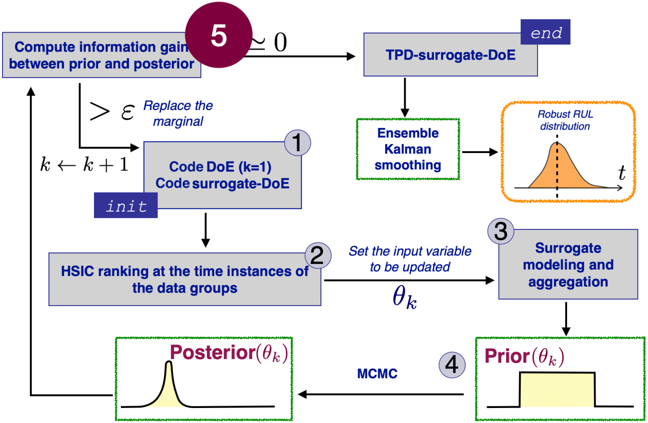
This ensures robustness by leveraging multiple models while minimizing bias

Step 4



4. Estimate the posterior distribution $\widehat{p}(\theta_k | \boldsymbol{y}^1, \ldots, \boldsymbol{y}^q, \boldsymbol{u}_{0,k})$ with an MCMC sampling procedure

Step (5)



5. Compute the Kullback-Leibler divergence d_{KL} between prior distribution $\mathcal{U}(\theta_k)$ and the estimated density:
   - If $d_{KL} > \varepsilon$, update the prior $\mu_{\mathbf{x},k}$ by replacing marginal $\mathcal{U}(\theta_k)$ with $\widehat{p}(\theta_k|\mathbf{y}^1, \ldots, \mathbf{y}^q, \mathbf{u}_{0,k})$ and continue $k \leftarrow k+1$
   - Otherwise, stop and obtain an *updated* RUL prediction by computing $g\#\mu_{\mathbf{x},k^*}$

Bayesian updating step 4

Proposition

Assume $\lambda := 1/\sigma_{\boldsymbol{\eta}}^2 \sim \mathcal{G}(\frac{m}{2}, \frac{1}{2}\|\mathbf{y} - f(\theta)\|^2)$ *(Gamma distribution), where* m *is the number of data points in* $\mathbf{y}$; $\theta \sim \mathcal{U}(\theta)$, *and* $p(\theta, \lambda) \propto \lambda^{-1}$. *Then:*

$$p(\theta|\mathbf{y}) \propto \|\mathbf{y} - f(\theta)\|^{-m} \tag{12}$$

Moreover, if multiple groups of data at different time-instances are considered, $\mathbf{y}^1, \ldots, \mathbf{y}^q$, *with respective priors on the inverse of their standard deviations* $\lambda_i \sim \mathcal{G}(\frac{m_i}{2}, \frac{1}{2}\|\mathbf{y}^i - f(\theta)\|^2)$, *then the generalization is:*

$$p(\theta|\mathbf{y}^1, \ldots, \mathbf{y}^q) \propto \prod_{i=1}^{q} \|\mathbf{y}^i - f(\theta)\|^{-m_i} \tag{13}$$

Proof. Bayes' theorem and simplifications.

Bayesian updating step 4

$$p(\theta_k|\mathbf{y}^1,\ldots,\mathbf{y}^q,\mathbf{u}_{0,k}) \propto \frac{1}{M}\sum_{r=1}^{M}\prod_{i=1}^{q}\|\mathbf{y}^i - \langle \mathbf{w}^{(r)}, \widehat{\mathbf{g}}(\mathbf{u}_{0,k},\theta_k)\rangle\|^{-m_i} \qquad (14)$$

- Use Random Walk Metropolis-Hastings (RWMH) MCMC algorithm [Sullivan, 2015] to sample from (14)
- Monte-Carlo integration using sample $\{\mathbf{w}^{(r)}\}_{r=1}^{M}$ from the Dirichlet-$\mathbb{1}_p$ distribution on the simplex $\rightarrow$ integrate hyperparameter
- Test convergence of RWMH chains with Gelman-Rubin test [Gelman and Rubin, 1992]
- Updated densities are conditioned on nominal values $\mathbf{u}_{0,k}$ of *other* $d-1$ *input variables* $\rightarrow$ future work on how to integrate uncertainty
- Use log-sum-exp trick for numerical computation

- We don't need sequential updating at each time-step $p(g(t_k, \boldsymbol{X})|\mathcal{D})$ (which is the goal of filtering), but to assimilate the data on the entire window frame
- Indeed the data is *unavailable* at each time-step, so we need to account of all the information on a window
- The good paradigm for this is smoothing, i.e getting estimates from the full-posterior distribution $p(g(t_1, \boldsymbol{X}), \dots, g(t_k, \boldsymbol{X})|\mathcal{D}) \rightarrow$ Ensemble Kalman smoothing (EnKS) [Evensen and van Leeuwen, 2000]

- Build an ensemble $\{g(\boldsymbol{X}^{(k)})\}_{k=1}^{n} \sim g\#h_{\boldsymbol{X}|\mathcal{D}}$, suppose for each data-group $i = 1, \ldots, q$:

$$y^i(t_\ell^i) = g(t_\ell^i, \boldsymbol{X}) + \eta_i, \tag{15}$$

where the variance R^i of the noise *is known*

- Define the ensemble mean and anomalies:

$$\overline{g}(t) = \frac{1}{n}\sum_{k=1}^{n} g(t, \boldsymbol{X}^{(k)}), \quad A^{(k)}(t) = g(t, \boldsymbol{X}^{(k)}) - \overline{g}(t) \tag{16}$$

- The cross-covariance between ensemble states at any time t and the observation times t_ℓ^i are approximated empirically:

$$\widehat{C}(t, t_\ell^i) = \frac{1}{n-1}\sum_{k=1}^{n} A^{(k)}(t) A^{(k)}(t_\ell^i) \tag{17}$$

- We define the Kalman gain for each t in the time-window:

$$K_t = \frac{\widehat{C}(t, t_\ell^i)}{\widehat{C}(t, t_\ell^i) + R^i} \tag{18}$$

- And defining the innovation term as $d_\ell^{(k)} = y^i(t_\ell^i) - g(t_\ell^i, \boldsymbol{X}^{(k)})$, we update all the ensembles following:

$$g(\boldsymbol{X}^{(k)}) \leftarrow g(\boldsymbol{X}^{(k)}) - K_t \cdot d_\ell^{(k)} \tag{19}$$

- This process is repeated *sequentially* across all observation times t_ℓ^i, for all data groups $i \in \{1, \ldots, q\}$

Outline

- Clogging of SGs is a complex multiphysics phenomenon that occurs following long operational periods in pressurized-water reactors (PWR) of the French nuclear fleet $\rightarrow$ undermines performance & weakens the structures $\rightarrow$ *may require chemical cleanings*

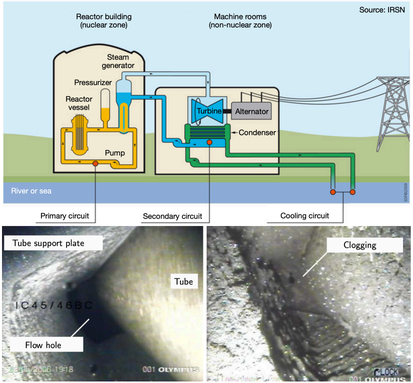
- 

Figure 2: PWR scheme, and example of video examination during a PWR outage (© IRSN, EDF)

- No state-of-the-art model allowing for ground insights on diagnostics and prognostics of clogging rate τ_c $\rightarrow$ very hard to model $\&$ challenging to create reproducible lab experiment for model validation $+$ not a lot of literature [Srikantiah and Chappidi, 2000; Prusek et al., 2013; Girard, 2014; Yang et al., 2017]
- Available scarce video field data as well as indirect measurements $\rightarrow$ allow to construct data-driven regression algorithms [Pinciroli et al., 2021] $\approx$ not enough data to have robust predictive models
- Another tool is the physical clogging model developed by [Prusek et al., 2013] $\rightarrow$ subsequent numerical model THYC-Puffer-DEPO [Feng et al., 2023] $\approx$ lack of enough trustworthy field data for precise V&V
- Necessary decision-making on chemical cleaning planning under uncertainty $\rightarrow$ ***how to make use of the available knowledge and models for achieving reliable predictions?***

- Computer code THYC-Puffer-DEPO, complex multiphysics [Jaber et al., 2024b] , chaining of 3 codes $\rightarrow$ allows to simulate SG clogging on entire lifespan of the asset integrating past chemical cleanings and predicting future τ_c levels
- Two data groups $q = 2$, corresponding to field data and regression data

Input variable	Distribution
α	$\mathcal{U}(100, 103)$
β	$\mathcal{U}(0.02, 0.025)$
ϵ_e	$\mathcal{U}(0.2, 0.5)$
ϵ_c	$\mathcal{U}(0.01, 0.3)$
d_p	$\mathcal{U}(0.5, 10.0) \times 10^{-6}$
$\Gamma_p(0)$	$\mathcal{U}(1.0, 8.0) \times 10^{-9}$
a_v	$\mathcal{U}(0, 15) \times 10^{-4}$

Table 1: Probabilistic modeling of uncertain input variables

- $L = 3$ time windows, 5 MCMC chains are launched for GR convergence test, use uniform proposal distributions
- Computing time around 40 min $\rightarrow 5/7$ distributions are informed by the data, distinct modes for a_v

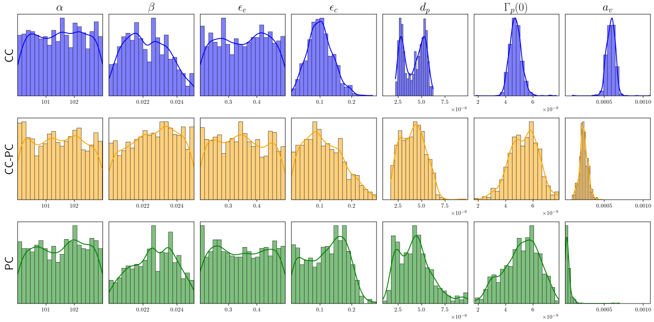


Figure 3: Posterior distributions of TPD clogging simulation code

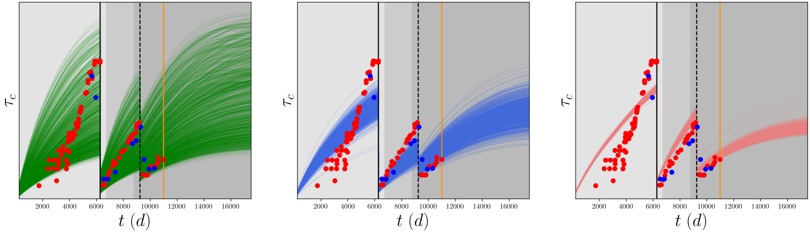
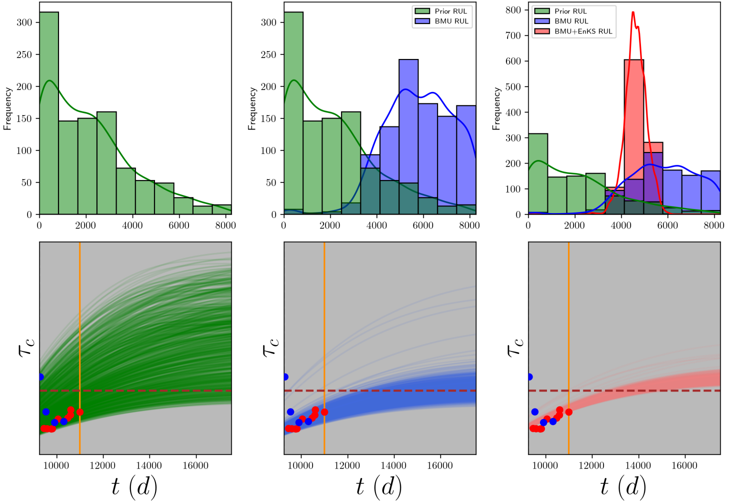


Figure 4: Prior/posterior and smoothed TPD emulations with Karhunen-Loève expansion surrogate

Results: posterior RUL



- RUL prediction uncertainty substantially reduced and mean of the distribution is shifted compared to the prior $\rightarrow$ positive impact for maintenance planning

Outline

Summary

Take-home messages:

- Offline data assimilation generic modular approach involving iterative BMU, non-intrusive emulators, heterogeneous data groups and ensemble Kalman smoothing
- Reduces parametric uncertainty of a computer simulation code that outputs future trajectories and RUL estimation, more informed decision making
- Demonstrated the approach on nuclear SG clogging use-case, showing reduced RUL uncertainty

Contributions:

- Paper published in **Reliability Engineering & System Safety** [Jaber et al., 2026]
- Reproducible code in `GitHub repository`

- How to integrate uncertainty in nominal parameters $\boldsymbol{u}_k$?
- Prior work on adaptive conformal prediction for GP surrogate models validation [Jaber et al., 2024a] (available GitHub repository)
- We define the cross-conformal estimator at a new point $\boldsymbol{X}^{(n+1)}$ using the posterior mean of the GP $\widetilde{g}$ and the posterior variance $\widetilde{\gamma}$:

$$\widehat{C}_{n,\alpha}^{J+GP}(\boldsymbol{X}^{(n+1)}) = \left[\widehat{q}_{n,\alpha}^{\pm}\left\{\widetilde{g}_{-i}(\boldsymbol{X}^{(n+1)}) \pm R_i^{LOO\gamma} \times \widetilde{\gamma}_{-i}(\boldsymbol{X}^{(n+1)})\right\}\right]$$

where $R_i^{LOO\gamma}$ is the Leave-One-Out error normalized by $\widetilde{\gamma}$

- Since the intervals are adaptive, can we use it as a proxy for metamodel accuracy for some step in the methodology? Surrogate refinement, or active learning?

Thank you for your attention!

Figure 5: Paper

Figure 6: GitHub

Let's connect: `edgar.jaber@ens-paris-saclay.fr`

Outline

- Hilbert-Schmidt Independence Criterion (HSIC) [Gretton et al., 2005], kernel method $\rightarrow$ evaluates sensitivity of a single-input in a given-data context, no need for surrogate models
- Theoretical result for all $i \in \{1, \ldots, d\}, k \in \{1, \ldots, N\}$:

$$\mathrm{HSIC}(X_i, g(\boldsymbol{X}, t_k)) = 0 \iff X_i \perp g(\boldsymbol{X}, t_k) \tag{20}$$

- The index disposes of U-stat and V-stat estimators $+$ hypothesis testing with corresponding p-value $\rightarrow$ implemented in the OpenTURNS
- The normalized R^2_{HSIC} index is better suited for interpretation:

$$R^2_{\mathrm{HSIC}}(X_i, g(\boldsymbol{X}, t_k)) = \frac{\mathrm{HSIC}(X_i, g(\boldsymbol{X}, t_k))}{\sqrt{(\mathrm{HSIC}(X_i, X_i)\mathrm{HSIC}(g(\boldsymbol{X}, t_k), g(\boldsymbol{X}, t_k)))}} \in [0, 1]$$

- Empirical evidence suggests that R^2_{HSIC} can be used confidently for variable ranking $\rightarrow$ HSIC-ANOVA decompositions also exist *but* only pathological cases create stark differences (see [Sarazin et al., 2022])

Evensen, G. and van Leeuwen, P. J. (2000). An ensemble Kalman smoother for nonlinear dynamics. *Monthly Weather Review*, 128:1852–1867.

Feng, Q., Nebes, J., Bachet, M., Pujet, S., You, D., and Deri, E. (2023). Tube support plates blockage of PWR steam generators: thermalhydraulics and chemical modeling. Proceedings of the International Conference on Water Chemistry in Nuclear Reactor Systems (NPC 2023), Juan-les-Pins.

Geir Evensen, Femke C. Vossepoel, P. J. v. L. (2022). *Data assimilation fundamentals*. Springer Cham.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457 – 472.

Girard, S. (2014). *Clogging of recirculating nuclear steam generators*. Springer International Publishing.

Gretton, A., Bousquet, O., Smola, A., and Schölkopf, B. (2005). Measuring statistical dependence with Hilbert-Schmidt norms. In *Proceedings Algorithmic Learning Theory*, pages 63–77. Springer-Verlag.

Jaber, E., Blot, V., Brunel, N., Chabridon, V., Remy, E., Iooss, B., Lucor, D., Mougeot, M., and Leite, A. (2024a). Conformal Approach To Gaussian Process Surrogate Evaluation With Coverage Guarantees.

Jaber, E., Chabridon, V., Remy, E., Baudin, M., Lucor, D., Mougeot, M., and Iooss, B. (2024b). Sensitivity analyses of a multi-physics long-term clogging model for steam generators. *International Journal for Uncertainty Quantification*.

Jaber, E., Remy, E., Chabridon, V., Mougeot, M., and Lucor, D. (2026). Fusion of heterogeneous data for robust degradation prognostics. *Reliability Engineering System Safety*, 274:112435.

Pinciroli, L., Baraldi, P., Shokry, A., Zio, E., Seraoui, R., and Mai, C. (2021). A semi-supervised method for the characterization of degradation of nuclear power plants steam generators. *Progress in Nuclear Energy*, 131:103580.

Prusek, T., Moleiro, E., Oukacine, F., Adobes, A., Jaeger, M., and Grandotto, M. (2013). Deposit models for tube support plate flow blockage in steam generators. *Nuclear Engineering and Design*, 262:418–428.

Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. The MIT Press, Massachusetts.

Sarazin, G., Marrel, A., Da Veiga, S., and Chabridon, V. (2022). Test d'indépendance basé sur les indices HSIC-ANOVA d'ordre total. Proceedings des 53èmes Journées de Statistique de la SFdS.

Srikantiah, G. and Chappidi, P. (2000). Particle deposition and fouling in PWR steam generators. *Nuclear Engineering and Design*, 200(1-2):285–294.

Sullivan, T. J. (2015). *Introduction to uncertainty quantification*. Texts in Applied Mathematics, 63. Springer-Verlag.

Vachtsevanos, G., Lewis, F., Roemer, M., Hess, A., and Wu, B. (2006). *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. John Wiley & Sons.

Yang, G., Pointeau, V., Tevissen, E., and Chagnes, A. (2017). A review on clogging of recirculating steam generators in pressurized-water reactors. *Progress in Nuclear Energy*, 97:182–196.