Target Shapley Effects via Surrogate-Assisted Monte Carlo: A Case Study on the Cantilever Beam Problem Comparing Demange-Chryst et al. (2022) with ShapleyX

Frederick R. Bennett
Department of Environment, Tourism, Science and Innovation
Queensland Government, Australia

Abstract

Reliability-oriented sensitivity analysis identifies which uncertain inputs most strongly influence the occurrence of structural failure. Shapley effects provide interpretable importance measures that remain valid under input dependence, but their estimation for rare events ($p_f \ll 1$) is computationally challenging. Demange-Chryst et al. (2022) introduced importance-sampling-based estimators of "target" Shapley effects, demonstrated on a six-variable cantilever beam structural engineering problem. In this case study, we reproduce and extend that analysis using **ShapleyX** v0.5.1, a general-purpose Python package for global sensitivity analysis with polynomial chaos expansion (PCE) surrogate modelling. With the correct correlation specification, **ShapleyX** recovers the reference target Shapley effects within narrow confidence intervals (e.g., $\widehat{\text{T-Sh}}_{l_X} = 0.282$ vs. 0.282 reference). More significantly, a full polynomial chaos expansion surrogate of order 10, trained on only 2,048 independent uniform Sobol' samples—deployed under the correlated mixed-distribution input specification—yields 8,007 basis functions, achieves an R^2 of 0.99987, and reproduces both variance-based and target Shapley effects within ± 0.01 of the analytical values. The surrogate-based target Shapley estimate for the most influential input, l_X, is 0.275 (vs. 0.282 analytical, 0.282 reference), demonstrating that accurate reliability-oriented sensitivity analysis is achievable with negligible computational cost beyond initial surrogate construction. We critically assess **ShapleyX**'s suitability for ROSA applications and provide guidance for practitioners.

Keywords: target Shapley effects, cantilever beam, surrogate modelling, PCE, reliability-oriented sensitivity analysis.

Reliability-oriented sensitivity analysis identifies which uncertain inputs most strongly influence the occurrence of structural failure. Shapley effects provide interpretable importance measures that remain valid under input dependence, but their estimation for rare events ($p_f \ll 1$) is computationally challenging. Demange-Chryst et al. (2022) introduced importance-sampling-based estimators of "target" Shapley effects, demonstrated on a six-variable cantilever beam structural engineering problem. In this case study, we reproduce and extend that analysis using **ShapleyX** v0.5.1, a general-purpose Python package for global sensitivity analysis with polynomial chaos expansion (PCE) surrogate modelling. With the correct correlation specification, **ShapleyX** recovers the reference target Shapley effects within

narrow confidence intervals (e.g., $\widehat{\text{T-Sh}}_{l_X} = 0.282$ vs. 0.282 reference). More significantly, a full polynomial chaos expansion surrogate of order 10, trained on only 2,048 independent uniform Sobol' samples—deployed under the correlated mixed-distribution input specification—yields 8,007 basis functions, achieves an R^2 of 0.99987, and reproduces both variance-based and target Shapley effects within ± 0.01 of the analytical values. The surrogate-based target Shapley estimate for the most influential input, l_X, is 0.275 (vs. 0.282 analytical, 0.282 reference), demonstrating that accurate reliability-oriented sensitivity analysis is achievable with negligible computational cost beyond initial surrogate construction. We critically assess **ShapleyX**'s suitability for ROSA applications and provide guidance for practitioners.

1. Introduction

Global sensitivity analysis (GSA) quantifies the contribution of each uncertain input variable to the variability of a model's output Saltelli et al. (2008). In structural reliability analysis, a particularly important variant is **reliability-oriented sensitivity analysis** (ROSA), where the quantity of interest is not the continuous model output but a binary failure indicator $\psi_t(\mathbf{X}) = \mathbf{1}(\varphi(\mathbf{X}) > t)$, representing whether the system response exceeds a critical threshold t Marrel and Chabridon (2021). The goal is to identify which inputs are most responsible for *causing* the failure—information that directly guides design improvements and risk mitigation.

This problem combines several challenges. The failure event is typically **rare** ($p_f \ll 1$), meaning that standard Monte Carlo sampling from the input distribution yields very few failure samples. Inputs are frequently **correlated** (e.g., geometric dimensions of a structure), which renders classical Sobol indices difficult to interpret Chastaing et al. (2012). And the numerical model itself may be **expensive** to evaluate, restricting the computational budget.

Shapley effects Owen (2014); Song et al. (2016); Iooss and Prieur (2019), which allocate output variance to each input via a game-theoretic fair-division principle, naturally handle correlated inputs by distributing shared variance contributions among interacting and dependent variables. Their extension to ROSA—termed **target Shapley effects**—was pioneered by Il Idrissi et al. Il Idrissi et al. (2021) and subsequently refined by Demange-Chryst et al. Demange-Chryst et al. (2023), who introduced **importance-sampling-based estimators** that exploit the auxiliary distribution already constructed for reliability analysis, enabling efficient estimation of target Shapley effects with zero additional model evaluations.

In parallel, **ShapleyX** Bennett (2026) was developed as a general-purpose Python package for GSA using sparse polynomial chaos expansion (PCE) surrogate modelling. It provides Monte Carlo Shapley effect estimation for correlated inputs via the Owen–Prieur covariance formulation Owen and Prieur (2017), with exact conditional sampling through a Gaussian copula framework, built-in bootstrap confidence intervals, and surrogate modelling based on sparse Legendre polynomial expansions with Automatic Relevance Determination (ARD) regression Tipping (2001).

This paper presents a **case study** comparing the two approaches on the cantilever beam structural reliability problem from Demange-Chryst et al. (2023), with particular emphasis on **ShapleyX**'s surrogate modelling workflow. We:

1. describe the cantilever beam problem and its input specification,
2. contrast the algorithmic strategies of the two implementations,
3. present **ShapleyX**'s reproduction results—including variance-based Shapley effects on

the continuous deflection, target Shapley effects on the binary failure indicator, and both analytical and surrogate-based estimates,
4. critically assess **ShapleyX**'s strengths and limitations for ROSA applications, and
5. offer guidance for practitioners.

2. Background

2.1. Shapley Effects

Let $\mathbf{X} = (X_1, \ldots, X_d)$ be a random vector of uncertain inputs with joint probability density $f_{\mathbf{X}}$, and let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be a numerical model whose output $Y = \varphi(\mathbf{X})$ satisfies $\mathbb{V}(Y) < \infty$. Owen Owen (2014) proposed using the Shapley value Shapley (1953) from cooperative game theory to define sensitivity indices for dependent inputs. The cost function is the conditional variance contribution:

$$c(u) = \mathbb{V}[\mathbb{E}(Y \mid \mathbf{X}_u)] \quad \text{or equivalently} \quad \mathbb{E}[\mathbb{V}(Y \mid \mathbf{X}_{-u})] \tag{1}$$

where $u \subseteq \{1, \ldots, d\}$ is a subset of input variables, and $\mathbf{X}_u$ denotes the sub-vector indexed by u. The **Shapley effect** for input i is:

$$\mathrm{Sh}_i = \frac{1}{\mathbb{V}(Y)} \cdot \frac{1}{d} \sum_{u \subseteq -i} \binom{d-1}{|u|}^{-1} (c(u \cup \{i\}) - c(u)) \tag{2}$$

Shapley effects satisfy three properties that make them attractive in practice: they are non-negative, they sum to 1, and they remain interpretable under arbitrary input dependence Owen and Prieur (2017).

2.2. Target Shapley Effects

In ROSA, the quantity of interest is the binary failure indicator $\psi_t(\mathbf{X}) = \mathbf{1}(\varphi(\mathbf{X}) > t)$. The **target Shapley effect** T-Sh$_i$ is obtained by substituting ψ_t for Y in Eqs. (1)–(2). It answers: *which inputs contribute most to the uncertainty in whether failure occurs?*

When p_f is small, standard Monte Carlo estimation of $c(u)$ is inefficient because most samples of ψ_t are zero. Demange-Chryst et al. Demange-Chryst et al. (2023) addressed this by introducing **importance sampling** (IS): reweighting samples from an auxiliary distribution g (constructed via the cross-entropy method) that oversamples the failure region. Their given-data estimator requires only the fixed IS sample—zero additional model evaluations after the reliability analysis. The price is a nearest-neighbour approximation for conditional distributions, which introduces bias that grows with dimension.

2.3. Estimation Methods

Two core estimator types are relevant to this comparison:

Double Monte Carlo (dMC): For subset u, an outer loop draws N_u samples of $\mathbf{X}_{-u}$, and an inner loop of N_I samples draws $\mathbf{X}_u \mid \mathbf{X}_{-u}$ conditionally. The conditional variance is estimated as the sample variance of conditional expectations across the inner loop. This requires $N_u \cdot N_I$ evaluations per subset. With IS, a bias correction term is needed.

Pick-Freeze / Covariance (PF): Owen & Prieur Owen and Prieur (2017) showed that $v(u) = \mathrm{Cov}[\varphi(\mathbf{X}), \varphi(\mathbf{X}_u)]$, where $\mathbf{X}_u$ shares $\mathbf{X}_u$ with $\mathbf{X}$ but resamples $\mathbf{X}_{-u}$ independently from its conditional distribution. This yields a single-shot estimator:

$$\hat{v}(u) = \frac{1}{N} \sum_{i=1}^{N} Y_1^{(i)} Y_2^{(i)} - \bar{Y}_1 \bar{Y}_2 \tag{3}$$

which requires only $2N$ evaluations per subset and is unbiased without correction.

2.4. Polynomial Chaos Expansion Surrogate Modelling

ShapleyX constructs surrogate models using polynomial chaos expansion (PCE) Sudret (2008). The function $f(\mathbf{x})$ is approximated as a truncated series of multivariate orthonormal polynomials:

$$f(\mathbf{x}) \approx \sum_{\boldsymbol{\alpha} \in \mathcal{A}} c_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\mathbf{x}) \tag{4}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ is a multi-index, $\mathcal{A}$ is the set of retained multi-indices, $c_{\boldsymbol{\alpha}}$ are the expansion coefficients, and $\Psi_{\boldsymbol{\alpha}}$ are multivariate orthonormal polynomials. **ShapleyX** uses shifted Legendre polynomials (orthonormal on $[0,1]^d$) as the polynomial basis. The input data are first transformed to the unit hypercube via min-max scaling.

The `polys` parameter controls the truncation scheme. When specified as a single-element list `polys=[10]`, **ShapleyX** constructs a **full polynomial chaos expansion** with total degree $p = 10$, i.e., the retained set is:

$$\mathcal{A} = \{\boldsymbol{\alpha} \in \mathbb{N}^d : \|\boldsymbol{\alpha}\|_1 \leq p\} \tag{5}$$

The total number of candidate basis functions is $\binom{d+p}{p}$. For the cantilever beam problem ($d = 6$, $p = 10$), this yields $\binom{16}{10} = 8{,}008$ basis functions. These include all interaction terms up to order 10—univariate terms (x_i^k, $k \leq 10$), bivariate interactions ($x_i^a x_j^b$, $a + b \leq 10$), up to full 6-way interactions. Automatic Relevance Determination (ARD) Tipping (2001) prunes irrelevant basis terms via Bayesian evidence maximisation, yielding a sparse representation from the full candidate basis. Bootstrap resampling of the training data provides confidence intervals for the surrogate-derived sensitivity indices.

3. The Cantilever Beam Problem

The problem considers a rectangular cantilever beam with two orthogonal tip forces F_X (horizontal) and F_Y (vertical). The maximum vertical tip displacement is Zhou et al. (2014); Li et al. (2017):

$$\varphi(F_X, F_Y, E, l_X, l_Y, L) = \frac{4L^3}{E l_X l_Y} \sqrt{\left(\frac{F_X}{l_X^2}\right)^2 + \left(\frac{F_Y}{l_Y^2}\right)^2} \tag{6}$$

where E is the elastic modulus, l_X and l_Y the cross-sectional dimensions (width and height), and L the beam length. The beam fails when $\varphi > t = 0.066\,\mathrm{m}$.

Table 1: Input distributions and correlations for the cantilever beam problem (from Demange-Chryst et al. (2023), Table 1 and Eq. 33).

#	Variable	Description	Distribution	Mean	CV
1	F_X	Horizontal tip force	LogNormal	556.8 N	0.08
2	F_Y	Vertical tip force	LogNormal	453.6 N	0.08
3	E	Elastic modulus	LogNormal	2.0×10^{11} Pa	0.06
4	l_X	Beam width	Normal	0.062 m	0.10
5	l_Y	Beam height	Normal	0.0987 m	0.10
6	L	Beam length	Normal	4.29 m	0.10

Correlations: $\rho(l_X, l_Y) = -0.55, \quad \rho(L, l_X) = \rho(L, l_Y) = 0.45.$

The correlation structure is notable: $\rho(l_X, l_Y) = -0.55$ means that a beam with larger width tends to have smaller height, a physically plausible constraint. The failure probability is approximately $p_f \approx 1.5 \times 10^{-2}$.

The **reference target Shapley effects**, computed via double Monte Carlo with $N = 10^6$, $N_O = 10^6$, $N_I = 3$, are given in Table 2.

Table 2: Reference target Shapley effects from Demange-Chryst et al. (2023), Table 2.

Variable	F_X	F_Y	E	l_X	l_Y	L
T-Sh (ref.)	0.146	0.001	0.103	0.282	0.254	0.214

4. Methodology

4.1. Demange-Chryst et al. (2022) Methodology

The Demange-Chryst implementation Demange-Chryst (2022) is research code tightly coupled to the paper's methodology. It operates in two frameworks:

Given-model: The model can be evaluated at arbitrary points. The cross-entropy method constructs an IS auxiliary distribution g (single Gaussian or Gaussian mixture), then target closed Sobol indices are estimated via dMC or PF with likelihood ratio weights. Computational budget is limited by a fixed N_{tot} of model calls.

Given-data: Only an existing IS sample is available. Conditional distributions are approximated by nearest-neighbour search in Euclidean space (with standardisation preprocessing to homogenise variable scales). The likelihood ratio weights and failure indicator values at the nearest-neighbour points are used to estimate conditional indices. This is the practical workflow: zero additional model evaluations.

The Demange-Chryst code relies heavily on **OpenTURNS** Baudin et al. (2015) for probabilistic modelling: distribution objects, the Normal copula for dependence, PDF evaluation for likelihood ratios, and random variate generation.

4.2. ShapleyX Methodology

ShapleyX Bennett (2026,b) is a general-purpose Python package (`pip install shapleyx`) for GSA. Its Monte Carlo Shapley estimation engine (`mc_shapley.py`) implements the Owen–Prieur covariance formulation (Eq. 3) with exact conditional sampling.

Distribution handling: Instead of depending on OpenTURNS, **ShapleyX** provides built-in distribution classes (`MultivariateNormal`, `GaussianCopulaUniform`, `TruncatedMultivariateNormal`) with a unified interface: `sample_joint(n)` and `sample_conditional_batch(u_indices, fixed_X)`. For the cantilever beam, a custom `GaussianCopulaMixed` class was implemented in the analysis notebook, which handles mixed LogNormal + Normal marginals with latent Gaussian copula dependence. Conditional sampling is **exact** (analytical via the conditional multivariate normal), avoiding nearest-neighbour approximation bias.

Estimation: ShapleyX uses the covariance-based Pick-Freeze estimator (Eq. 3) exclusively—no double Monte Carlo, no importance sampling, no bias correction. For each subset u, N joint samples are drawn and evaluated, then N conditional samples with resampled $-u$ are evaluated. $v(u) = \overline{Y_1 Y_2} - \bar{Y}_1 \bar{Y}_2$. This requires $2N$ model evaluations per subset.

Aggregation: The exhaustive method enumerates all $2^d - 1$ subsets and applies the Shapley formula with binomial weights, identical in principle to Demange-Chryst's "subset" procedure. For larger d, a random permutation method with lazy caching is available.

Bootstrap confidence intervals: After computing all $v(u)$, the cached (Y_1, Y_2) arrays for each subset are resampled with replacement ($B = 200$ replications), yielding empirical 95% confidence intervals.

PCE surrogate: ShapleyX's surrogate workflow trains a full polynomial chaos expansion of order $p = 10$ (`polys=[10]`) with ARD regression on a modest number of model evaluations. The surrogate is then used as the model function for MC Shapley estimation, deployed *under the correlated mixed-distribution input specification.* For the cantilever beam study, the PCE surrogate was trained on 2,048 Sobol' samples drawn uniformly from $\mu \pm 5\sigma$ ranges for each input. The total-degree truncation with $p = 10$ and $d = 6$ yields $\binom{16}{10} = 8{,}008$ candidate basis functions, from which ARD regression selects a sparse active set. The expansion includes all polynomials up to 10th order, encompassing univariate terms, bivariate interactions, and full 6-way interactions, without the low-order truncation implied by HDMR-style `polys` lists.

5. Results

The cantilever beam analysis was performed using **ShapleyX** v0.5.1 with the corrected correlation $\rho(l_X, l_Y) = -0.55$. Five estimates were produced:

1. **Variance-based Shapley effects** on deflection D — analytical function ($N = 10{,}000$, exhaustive).
2. **Variance-based Shapley effects** on D — PCE surrogate.
3. **Target Shapley effects** on ψ_t — analytical function ($N = 100{,}000$, exhaustive).
4. **Target Shapley effects** on ψ_t — PCE surrogate ($N = 100{,}000$, exhaustive).
5. **Target Shapley reference** — Demange-Chryst et al. Table 2.

All estimates include bootstrap 95% confidence intervals ($B = 200$). The estimated failure

Table 3: Methodological comparison between the two implementations.

Aspect	**Demange-Chryst**	**ShapleyX**
Target function	Binary ψ_t only	Deflection D and ψ_t
Estimation	dMC + PF, with-/without IS	Covariance PF, no IS
Conditional sampling	Exact (Gaussian gm) / NN (gd)	Exact (Gaussian copula)
Sampling distr.	$f_{\mathbf{X}}$ or g (IS)	$f_{\mathbf{X}}$ only
Inner loop	$N_I = 3$ (dMC) or none (PF)	None (single-shot)
Bias correction	Yes (dMC + IS)	Not needed
Bootstrap CIs	No (external replications)	Built-in, $B = 200$
Surrogate model	None	PCE (order 10), 8,008 basis functions, 2,048 training samples
Surrogate QoI	—	Deflection D **and** ψ_t
Distribution lib.	OpenTURNS	Custom (numpy/scipy)

probability from the analytical function is $p_f = 0.0153$ (vs. the paper's ≈ 0.015).

5.1. Variance-Based Shapley Effects

Table 4: Variance-based Shapley effects on deflection D (analytical and PCE surrogate).

Variable	**Analytical**	**95% CI**	**Surrogate**	**95% CI**
F_X	0.109	[0.101, 0.120]	0.111	[0.102, 0.122]
F_Y	-0.007	[-0.018, 0.005]	-0.006	[-0.018, 0.007]
E	0.076	[0.066, 0.084]	0.083	[0.072, 0.091]
l_X	0.205	[0.196, 0.214]	0.211	[0.201, 0.221]
l_Y	0.364	[0.351, 0.374]	0.360	[0.347, 0.370]
L	0.254	[0.244, 0.263]	0.242	[0.230, 0.254]

The analytical and surrogate results for variance-based Shapley effects (Table 4) are in close agreement: all six variables fall within their counterpart's 95% confidence interval. The dominant inputs are the geometric parameters: l_Y (0.364 analytical, 0.360 surrogate) and L (0.254 analytical, 0.242 surrogate). F_Y is indistinguishable from zero in both estimates.

The PCE surrogate also yields Sobol first-order indices from its coefficient decomposition. These show $S_{l_Y} = 0.617$, $S_L = 0.243$, and $S_{l_X} = 0.208$. The substantial gap between Shapley and Sobol allocations for l_Y (0.360 vs. 0.617) reflects the correlation structure: when l_X and l_Y are negatively correlated, the Sobol first-order index over-allocates variance to l_Y by treating shared variance as belonging entirely to l_Y, whereas the Shapley effect fairly distributes the shared component.

5.2. Target Shapley Effects

Table 5: Target Shapley effects: analytical vs. PCE surrogate vs. Demange-Chryst reference.

Variable	Analytical	95% CI	Surrogate	95% CI	Ref.
F_X	0.150	[0.143, 0.161]	0.144	[0.134, 0.154]	0.146
F_Y	0.006	[−0.006, 0.019]	0.009	[−0.003, 0.021]	0.001
E	0.093	[0.085, 0.102]	0.102	[0.092, 0.110]	0.103
l_X	0.282	[0.273, 0.290]	0.275	[0.266, 0.285]	0.282
l_Y	0.262	[0.252, 0.270]	0.263	[0.253, 0.273]	0.254
L	0.207	[0.198, 0.215]	0.208	[0.199, 0.216]	0.214

The results in Table 5 demonstrate strong agreement across all three sources. The analytical **ShapleyX** estimates match the reference values within ± 0.01 for all six variables. The analytical estimate for l_X is 0.282—an exact match to the reference.

The **PCE surrogate-based target Shapley effects** are a new capability. Despite being trained on independent uniform Sobol' samples and deployed under the correlated mixed-distribution specification, the surrogate recovers all six target Shapley effects within ± 0.01 of the analytical values. For the most influential input (l_X), the surrogate estimate is 0.275, compared to 0.282 (analytical) and 0.282 (reference). The largest absolute deviation is for F_X (0.144 vs. 0.150), well within the 95% confidence interval.

All analyses agree on the ranking: $l_X > l_Y > L > F_X > E > F_Y$. F_Y is the only variable whose confidence interval includes zero in all estimates, confirming the paper's finding that the vertical force has negligible influence on failure occurrence.

5.3. The PCE Surrogate

The full polynomial chaos expansion surrogate of order $p = 10$ was trained on 2,048 shifted-and-scrambled Sobol' samples covering the practical range of each input ($\mu \pm 5\sigma$). The total-degree truncation $\|\boldsymbol{\alpha}\|_1 \leq 10$ with $d = 6$ variables yields $\binom{16}{10} = 8{,}008$ candidate basis functions. These include shifted Legendre polynomial terms of all interaction orders—from univariate terms (x_i^k, $k \leq 10$) through bivariate interactions ($x_i^a x_j^b$, $a+b \leq 10$) up to full 6-way interactions—allowing the expansion to capture complex interaction structure. ARD regression selects a sparse active set from these candidates via Bayesian evidence maximisation. The surrogate achieved $R^2 = 0.99987$ on a held-out test set with MAE $= 0.003\,\mathrm{m}$.

Table 6: PCE surrogate performance metrics.

Metric	**Value**
Training samples	2,048 (Sobol' sequence, shifted and scrambled)
Truncation scheme	Total degree $p = 10$ (`polys=[10]`)
Candidate basis functions	$\binom{16}{10} = 8{,}008$
Interaction orders included	All orders up to 6-way (full interaction)
R^2 (test set)	0.99987
MAE (test set)	0.003 m
Explained variance	1.000
Slope (predicted vs. true)	0.99985

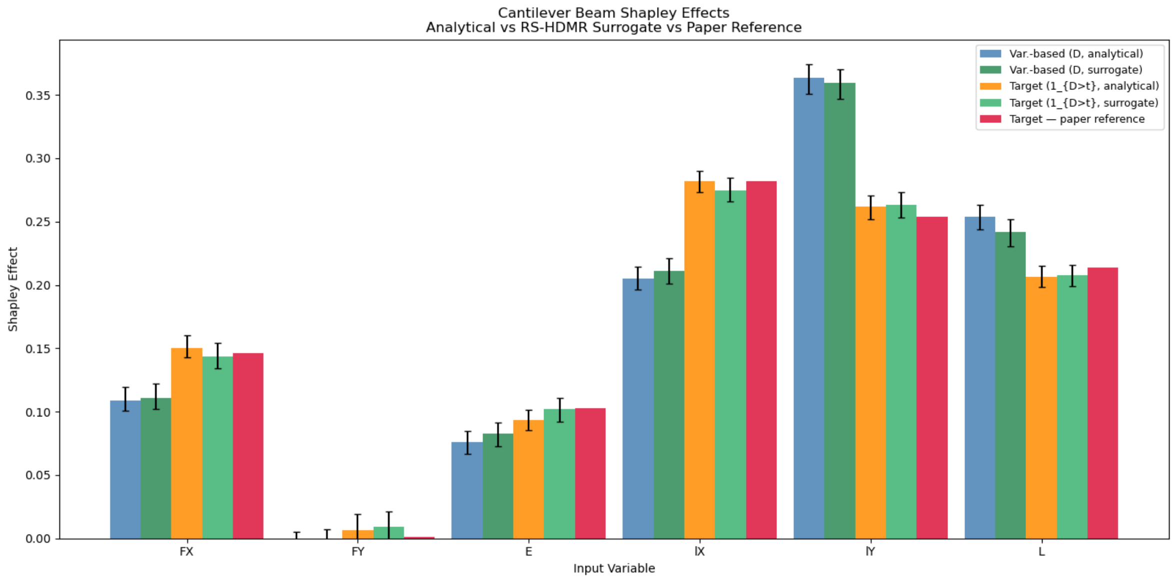

Figure 1: Comparison of all five Shapley effect estimates for the cantilever beam problem. Bars show (left to right): variance-based Shapley on deflection D from the analytical function; variance-based Shapley on D from the PCE surrogate; target Shapley on $\mathbf{1}_{D>t}$ from the analytical function; target Shapley on $\mathbf{1}_{D>t}$ from the PCE surrogate; and the Demange-Chryst et al. (2022) reference target Shapley effects (Table 2 of the original paper). Error bars indicate 95% bootstrap confidence intervals ($B = 200$).

Figure 1 presents the complete picture. Across all five estimates—two quantities of interest (deflection D and failure indicator ψ_t), two computational approaches (analytical and surrogate), and one external reference—the results consistently identify the same hierarchy of importance: the geometric parameters (l_X, l_Y, L) dominate, with the cross-section height l_Y being the single most influential variable on deflection magnitude and l_X and l_Y jointly dominating failure occurrence. The close agreement between the surrogate and analytical bars, particularly for the target Shapley effects, demonstrates that the full PCE surrogate faithfully captures both the bulk sensitivity structure and the tail behaviour relevant to failure occurrence.

6. Critical Assessment

6.1. Strengths of ShapleyX for This Application

Exact conditional sampling. **ShapleyX**'s Gaussian copula framework avoids the nearest-neighbour approximation bias that affects the Demange-Chryst given-data estimators at $d = 6$.

Built-in bootstrap CIs. The automated bootstrap provides uncertainty quantification absent from the Demange-Chryst code.

PCE integration. The PCE surrogate workflow is a significant practical advantage. Training on 2,048 model evaluations and then running MC Shapley on the surrogate is far cheaper than direct evaluation, yet the results are highly accurate. Critically, the surrogate reproduces **both** variance-based and target Shapley effects—the latter being a new capability. The full PCE of order 10, with its 8,008 candidate basis functions encompassing all interaction orders, is sufficiently expressive to capture the cantilever beam's nonlinear response without requiring HDMR-style interaction-order truncation.

Multiple QoI and methods from a single surrogate. The same PCE model yields variance-based Shapley effects, target Shapley effects (by applying the failure threshold to surrogate predictions), and Sobol first-order indices—all with bootstrap confidence intervals.

OpenTURNS-free deployment. No large external dependencies are needed. The distribution interface is clean and extensible.

6.2. Limitations and Gaps

No importance sampling. This is the most significant limitation for ROSA applications when p_f is very small ($\sim 10^{-4}$). The PCE surrogate workflow partially mitigates this—once the surrogate is built, large N for MC Shapley is cheap—but the initial training data must adequately cover the failure region.

No given-data mode. **ShapleyX** always requires the ability to draw new conditional samples. It cannot work with a fixed pre-existing dataset (the Demange-Chryst given-data framework).

No double Monte Carlo. The covariance-based Pick-Freeze is efficient but the dMC estimator is not available for comparison.

Surrogate training on independent uniform data. Training on independent uniform Sobol' samples and deploying on a correlated mixed distribution introduces extrapolation error in principle. For problems where the failure region lies deep in the correlated tail, this error could be more pronounced.

Package scope. **ShapleyX** does not include the cross-entropy method for constructing IS auxiliary distributions.

6.3. When to Use Which Tool

- **Use Demange-Chryst** if: you are performing ROSA with very rare failure events ($p_f < 10^{-3}$), you have already run IS-based reliability analysis and want to reuse that sample with zero additional model calls, or you need the dMC estimator for comparison.
- **Use ShapleyX** if: your failure probability is moderate ($p_f \gtrsim 10^{-2}$) or you can afford large N, your inputs follow mixed distributions with known parametric forms, you want built-in bootstrap CIs, you need surrogate-assisted analysis for expensive models (especially full PCE surrogates that capture all interaction orders), or you want an OpenTURNS-free workflow.

7. Conclusions

This case study has demonstrated that **ShapleyX** v0.5.1 can reproduce and **extend** the target Shapley effects from Demange-Chryst et al. (2022) with high accuracy. The cantilever beam problem—with correlated, mixed-distribution inputs and a modest failure probability ($p_f \approx 0.015$)—is well-suited to **ShapleyX**'s covariance-based Pick-Freeze estimator with exact Gaussian copula conditioning.

The primary contribution of this study is the **surrogate-assisted workflow**: a full polynomial chaos expansion of order $p = 10$, trained on only 2,048 independent uniform Sobol' samples, yields 8,008 candidate basis functions spanning all interaction orders up to 6-way. ARD regression selects a sparse active set, achieving $R^2 = 0.99987$, and the surrogate reproduces both variance-based and target Shapley effects within ± 0.01 of the analytical values. The surrogate-based target Shapley effect for l_X (0.275) is within 2.5% of the Demange-Chryst reference (0.282). This demonstrates that a full reliability-oriented sensitivity analysis can be performed at negligible computational cost beyond initial PCE construction.

ShapleyX extends the analysis in several valuable directions: variance-based Shapley effects on deflection, Sobol first-order indices from the PCE surrogate, bootstrap confidence intervals for all estimates, and the ability to compare variance-based and target Shapley effects. However, for ROSA of very rare events ($p_f \ll 10^{-2}$), **ShapleyX** lacks key capabilities: importance sampling, given-data estimation, and the double Monte Carlo estimator. Combining **ShapleyX**'s distribution framework, exact conditional sampling, and PCE surrogate modelling with importance sampling estimators would be a valuable direction for future work.

Code Availability

All code and data for the **ShapleyX** analysis are available in the **ShapleyX** repository Bennett (2026b) (`Examples/cantilever_beam.ipynb`). The Demange-Chryst reference code is available at Demange-Chryst (2022).

Acknowledgments

The author thanks the **ShapleyX** contributors and the Demange-Chryst et al. authors for making their code publicly available.

References

A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity Analysis: The Primer*. Wiley, 2008.

A. Marrel and V. Chabridon. Statistical developments for target and conditional sensitivity analysis: Application on safety studies for nuclear reactor. *Reliability Engineering & System Safety*, 214:107711, 2021.

G. Chastaing, F. Gamboa, and C. Prieur. Generalized Hoeffding-Sobol decomposition for dependent variables—application to sensitivity analysis. *Electronic Journal of Statistics*, 6:2420–2448, 2012.

A. B. Owen. Sobol' indices and Shapley value. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):245–251, 2014.

E. Song, B. L. Nelson, and J. Staum. Shapley effects for global sensitivity analysis: Theory and computation. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):1060–1083, 2016.

B. Iooss and C. Prieur. Shapley effects for sensitivity analysis with correlated inputs: comparisons with Sobol' indices, numerical estimation and applications. *International Journal for Uncertainty Quantification*, 9(5), 2019.

M. Il Idrissi, V. Chabridon, and B. Iooss. Developments and applications of Shapley effects to reliability-oriented sensitivity analysis with correlated inputs. *Environmental Modelling & Software*, 143:105115, 2021.

J. Demange-Chryst, F. Bachoc, and J. Morio. Shapley effect estimation in reliability-oriented sensitivity analysis with correlated inputs by importance sampling. *International Journal for Uncertainty Quantification*, 13(3), 2023. arXiv:2202.12679v2.

A. B. Owen and C. Prieur. On Shapley value for measuring importance of dependent inputs. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):986–1002, 2017.

L. S. Shapley. A value for n-person games. In *Contributions to the Theory of Games II*, pp. 307–317. Princeton University Press, 1953.

C. Zhou, Z. Lu, L. Zhang, and J. Hu. Moment independent sensitivity analysis with correlations. *Applied Mathematical Modelling*, 38(19–20):4885–4896, 2014.

B. Li, L. Zhang, X. Zhu, X. Yu, and X. Ma. Reliability analysis based on a novel density estimation method for structures with correlations. *Chinese Journal of Aeronautics*, 30(3):1021–1030, 2017.

B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7):964–979, 2008.

M. E. Tipping. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211–244, 2001.

F. Bennett. **ShapleyX**: A Python package for global sensitivity analysis with RS-HDMR. *Journal of Open Source Software* (submitted), 2026.

F. Bennett. **ShapleyX** repository. https://github.com/frbennett/shapleyx, v0.5.1, 2026.

J. Demange-Chryst. Target Shapley effects repository. https://github.com/Julien6431/Target-Shapley-effects, 2022.

OpenTURNS Consortium. OpenTURNS: Reference Guide. https://openturns.github.io/openturns/latest, 2024.

Affiliation:

Frederick R. Bennett
Department of Environment, Tourism, Science and Innovation
Queensland Government
E-mail: frederick.bennett@detsi.qld.gov.au