Shapley Effects for Rare-Event Sensitivity Analysis: A Case Study on the Rothermel Fire Spread Model 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 (ROSA) identifies which uncertain inputs most strongly influence the occurrence of rare failure events, but the combination of correlated inputs, mixed distribution types, and extremely low failure probabilities ($p_f \sim 10^{-4}$) makes estimation challenging. Demange-Chryst et al. (2022) introduced importance-sampling-based estimators of target Shapley effects specifically to address the rare-event regime, demonstrated on a 10-variable wildland fire spread model using the semi-physical Rothermel equations. In this case study, we apply **ShapleyX** v0.5.1—a general-purpose Python package for global sensitivity analysis—to the same problem, intentionally *without* importance sampling, as a stress test of its standard Monte Carlo Shapley estimation under rare-event conditions. Despite $p_f \approx 1.28 \times 10^{-4}$ meaning that only ~ 3 failures are expected per 20,000 samples, **ShapleyX**'s covariance-based Pick-Freeze estimator with $N = 300{,}000$ and exact Gaussian copula conditional sampling recovers the five most influential variables identified by the Demange-Chryst IS reference (σ, U, δ, m_l, m_d), with target Shapley effects generally within ± 0.05 of the reference values. The fuel particle area-to-volume ratio σ is identified as most important (T-Sh = 0.202 vs. 0.247 reference), followed by wind speed U (0.167 vs. 0.182) and fuel depth δ (0.157 vs. 0.152). The wider confidence intervals and larger discrepancies for low-influence variables (h, ρ_p, S_T, $\tan\phi$) quantify the penalty of forgoing importance sampling. We critically assess **ShapleyX**'s suitability for ROSA applications and provide guidance for practitioners.

Keywords: target Shapley effects, fire spread, Rothermel, rare events, surrogate modelling, ROSA.

Reliability-oriented sensitivity analysis (ROSA) identifies which uncertain inputs most strongly influence the occurrence of rare failure events, but the combination of correlated inputs, mixed distribution types, and extremely low failure probabilities ($p_f \sim 10^{-4}$) makes estimation challenging. Demange-Chryst et al. (2022) introduced importance-sampling-based estimators of target Shapley effects specifically to address the rare-event regime, demonstrated on a 10-variable wildland fire spread model using the semi-physical Rothermel equations. In this case study, we apply **ShapleyX** v0.5.1—a general-purpose Python package for global sensitivity analysis—to the same problem, intentionally *without* importance sampling, as a stress test of its standard Monte Carlo Shapley estimation under rare-event conditions. Despite $p_f \approx 1.28 \times 10^{-4}$ meaning that only ~ 3 failures are expected per 20,000 samples, **Shap-**

leyX's covariance-based Pick-Freeze estimator with $N = 300{,}000$ and exact Gaussian copula conditional sampling recovers the five most influential variables identified by the Demange-Chryst IS reference (σ, U, δ, m_l, m_d), with target Shapley effects generally within ± 0.05 of the reference values. The fuel particle area-to-volume ratio σ is identified as most important (T-Sh = 0.202 vs. 0.247 reference), followed by wind speed U (0.167 vs. 0.182) and fuel depth δ (0.157 vs. 0.152). The wider confidence intervals and larger discrepancies for low-influence variables (h, ρ_p, S_T, $\tan\phi$) quantify the penalty of forgoing importance sampling. The `GaussianCopulaFire` distribution class handles the complex specification—LogNormal, scaled LogNormal, Normal marginals with truncation rules and a strong $\rho(m_d, U) = -0.8$ correlation—with full acceptance rate and empirical correlation $\hat{\rho}(m_d, U) = -0.7987$. We critically assess **ShapleyX**'s current capabilities for ROSA with rare events and identify importance sampling as the most important missing capability for this application domain.

1. Introduction

Global sensitivity analysis (GSA) of rare-event models combines three challenging features: the quantity of interest is a binary failure indicator that is nearly always zero, the inputs are frequently correlated, and the model may involve complex distribution families with truncation rules. Shapley effects Owen (2014); Song et al. (2016); Iooss and Prieur (2019)—game-theoretic importance measures that naturally handle input dependence—have been extended to this reliability-oriented setting as **target Shapley effects** Il Idrissi et al. (2021); Demange-Chryst et al. (2023). However, estimating target Shapley effects when the failure probability p_f is on the order of 10^{-4} is computationally formidable: standard Monte Carlo sees failure samples so rarely that the conditional variance contributions become dominated by noise.

Demange-Chryst et al. Demange-Chryst et al. (2023) addressed this by developing **importance-sampling-based estimators** of target Shapley effects, demonstrated on the Rothermel wildland fire spread model with 10 correlated, mixed-distribution inputs and $p_f \approx 1.4 \times 10^{-4}$. Their results (Table 4 of the original paper) serve as the reference against which alternative estimation strategies can be measured.

ShapleyX Bennett (2026,b) is a general-purpose Python package for GSA using Monte Carlo Shapley estimation with exact Gaussian copula conditional sampling and built-in bootstrap confidence intervals. It does not currently implement importance sampling, making the fire spread problem a **stress test**: how much information can standard MC Shapley estimation extract under extreme rare-event conditions?

This case study presents that stress test. We:

1. describe the Rothermel fire spread model and its complex input specification,
2. implement a custom `GaussianCopulaFire` distribution class in **ShapleyX**,
3. compute target Shapley effects via the covariance-based Pick-Freeze estimator with $N = 300{,}000$,
4. compare against the Demange-Chryst IS reference values, and
5. critically assess **ShapleyX**'s capabilities and limitations for ROSA with rare events.

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 $\mathbb{V}(Y) < \infty$. The cost function is the conditional variance contribution Owen (2014):

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

where $u \subseteq \{1, \ldots, d\}$ is a subset of input variables. 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 are non-negative, sum to 1, and 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 substitutes ψ_t for Y in Eqs. (1)–(2). It answers: which inputs contribute most to the uncertainty in whether failure occurs?

When $p_f \ll 1$, standard MC estimation of $c(u)$ is inefficient because the vast majority of samples of ψ_t are zero. The Demange-Chryst et al. Demange-Chryst et al. (2023) given-data estimator uses an importance sampling (IS) auxiliary distribution g that over-samples the failure region, requiring zero additional model evaluations beyond the IS sample. Their reference target Shapley effects for the fire spread model (Table 4 of Demange-Chryst et al. (2023)) were computed with IS at $N = 10^7$, $N_O = 10^6$, $N_I = 3$.

2.3. Pick-Freeze Covariance Estimator

ShapleyX uses the covariance-based Pick-Freeze (PF) estimator Owen and Prieur (2017). For each subset u, the estimator is:

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

where $Y_1^{(i)} = \varphi(\mathbf{X}^{(i)})$ and $Y_2^{(i)} = \varphi(\mathbf{X}_u^{(i)})$ with $\mathbf{X}_u^{(i)}$ sharing the u-components with $\mathbf{X}^{(i)}$ but resampling $-u$ independently from $P(\mathbf{X}_{-u} \mid \mathbf{X}_u)$. This requires $2N$ evaluations per subset and is unbiased without correction. However, it does not exploit importance sampling—every evaluation is drawn from $f_{\mathbf{X}}$, and when $p_f \approx 10^{-4}$, the Paasche of failures in the sample is extremely small.

3. The Rothermel Fire Spread Model

3.1. Model Description

The Rothermel model Rothermel (1972) is a semi-physical model for wildland surface fire spread, with modifications by Albini Albini (1976) and Catchpole & Catchpole Catchpole and Catchpole (1991). The model was adapted for sensitivity analysis by Song et al. Song et al. (2016) and used by Demange-Chryst et al. as their Example 4.3 Demange-Chryst et al. (2023).

The rate of fire spread R (cm·s^{-1}) is computed from 10 fuel, weather, and terrain parameters through a chain of empirical equations. The model involves internal unit conversions between metric and imperial systems. The full computational procedure—fuel loading, unit conversion, Rothermel reaction intensity, propagating flux ratio, wind and slope factors, and effective heating number—is implemented in the companion notebook `Examples/fire_spread.ipynb` in the **ShapleyX** repository. Failure is defined as $R > t = 60$ cm·s^{-1}.

3.2. Input Specification

The 10 input variables have mixed Normal, LogNormal, and scaled LogNormal marginal distributions with several truncation rules and one strong correlation. Table 1 summarises the specification from Demange-Chryst et al. (Table 3 and Eq. 36 of Demange-Chryst et al. (2023)).

Table 1: Input variables for the Rothermel fire spread model. LogNormal parameters (μ, σ) refer to the underlying Normal: $\log X \sim \mathcal{N}(\mu, \sigma^2)$. Normal parameters are mean and standard deviation.

#	Symbol	Description	Distribution	μ	σ
1	δ	Fuel depth (cm)	LogNormal	2.19	0.517
2	σ	Fuel particle area-to-volume ratio (cm^{-1})	LogNormal, ≥ 5	3.31	0.294
3	h	Fuel particle low heat content (Kcal·kg^{-1})	LogNormal	8.48	0.063
4	ρ_p	Oven-dry particle density (g·cm^{-3})	LogNormal	−0.592	0.219
5	m_l	Moisture content of live fuel (g·g^{-1})	Normal, ≥ 0	1.18	0.377
6	m_d	Moisture content of dead fuel (g·g^{-1})	Normal	0.19	0.047
7	S_T	Fuel particle total mineral content (g·g^{-1})	Normal, ≥ 0	0.049	0.011
8	U	Wind speed at midflame height (km·h^{-1})	$6.9 \times$ LogNormal	1.0174	0.5569
9	$\tan\phi$	Slope tangent	Normal, ≥ 0	0.38	0.186
10	P	Dead fuel loading / total fuel loading	LogNormal, ≤ 1	−2.19	0.640

Correlation: $\rho(m_d, U) = -0.8$ (higher wind speeds are associated with lower dead fuel moisture).

Rejection rules: All negative values rejected; $S_T \leq 1$; $P \leq 1$; $\sigma \geq 3/0.6 = 5$.

The reference failure probability is $p_f \approx 1.4 \times 10^{-4}$ Demange-Chryst et al. (2023). Our empirical estimate from 500,000 samples is $p_f = 1.28 \times 10^{-4}$.

3.3. Reference Target Shapley Effects

Demange-Chryst et al. Demange-Chryst et al. (2023) provide reference target Shapley effects in their Table 4, reproduced here in Table 2. These were computed using IS with $N = 10^7$, $N_O = 10^6$, $N_I = 3$, and represent the gold standard against which we compare.

Table 2: Reference target Shapley effects from Demange-Chryst et al. (2022), Table 4. Computed with importance sampling ($N = 10^7$, $N_O = 10^6$, $N_I = 3$).

Variable	δ	σ	h	ρ_p	m_l	m_d	S_T	U	$\tan\phi$	P
T-Sh (IS ref.)	0.152	0.247	0.011	0.003	0.162	0.145	0.016	0.182	0.009	0.073

The paper identifies the five most influential variables as δ, σ, m_l, m_d, and U—the fuel and weather parameters that directly control the rate of spread physics.

4. Methodology

4.1. ShapleyX Distribution Handling

ShapleyX provides a Gaussian copula framework for handling correlated mixed-distribution inputs. For the fire spread model, we implement a custom `GaussianCopulaFire` class that:

- Stores marginals as (distribution_type, μ, σ) triples,
- Constructs a latent multivariate Normal with correlation matrix encoding $\rho(m_d, U) = -0.8$ and identity elsewhere,
- Implements `sample_joint(n)`: draws n samples in latent Normal space, maps to original space via inverse CDFs, and applies rejection rules (all negative values, $S_T \leq 1$, $P \leq 1$, $\sigma \geq 5$),
- Implements `sample_conditional_batch(u_indices, fixed_X)`: exact analytical conditioning via the conditional multivariate Normal in latent space, with inverse CDF mapping to original space and the same rejection rules.

The class handles three marginal types: `lognormal` (LogNormal with log-parameters), `lognormal_scaled` (scale $\times$ LogNormal, for wind speed $U = 6.9 \times \text{LogNormal}$), and `normal` (Normal with mean and standard deviation). The implementation passes validation: the empirical correlation $\hat{\rho}(m_d, U) = -0.7987$ (target: -0.8), and all marginal moments match their theoretical values.

4.2. MC Shapley Estimation

Target Shapley effects are estimated via the covariance-based Pick-Freeze estimator (Eq. 3) with:

- $N = 300{,}000$ samples per subset,
- Exhaustive subset enumeration ($2^{10} - 1 = 1{,}023$ subsets),
- Total evaluations: $2 \times 300{,}000 \times 1{,}023 = 613.5$ million (16.4 minutes on a single core),
- Bootstrap with $B = 200$ replications for 95% confidence intervals.

The target function is the binary failure indicator $\psi_t(\mathbf{x}) = \mathbf{1}(\texttt{Rothermel}(\mathbf{x}) > 60)$. With $p_f \approx 1.28 \times 10^{-4}$, each batch of 300,000 samples contains approximately 38 failure events.

This is the fundamental challenge: the PF estimator relies on the covariance between Y_1 and Y_2, and when both are almost always zero, the signal-to-noise ratio is extremely low.

4.3. No Surrogate Modelling

Unlike the cantilever beam and wing weight case studies, **no surrogate model is constructed** for this problem. The Rothermel model involves strongly nonlinear threshold behaviour (the failure indicator is binary), unit conversions, and exponential saturation terms that make it poorly suited to polynomial chaos expansion surrogates of limited fidelity. The Rothermel function itself is cheap to evaluate (analytical), so the computational bottleneck is purely the MC estimation cost, not model evaluation cost. This makes the fire spread case study a pure test of **ShapleyX**'s distribution handling and MC estimation capabilities.

4.4. Methodological Comparison

Table 3: Methodological comparison between the two implementations.

Aspect	**Demange-Chryst**	**ShapleyX**
Target function	Binary ψ_t only	Binary ψ_t
Estimation	dMC + PF, with IS	Covariance PF, no IS
Conditional sampling	Exact (gm) / NN approx. (gd)	Exact (Gaussian copula)
Sampling distribution	$f_{\mathbf{X}}$ or g (IS)	$f_{\mathbf{X}}$ only
Inner loop	$N_I = 3$ (dMC)	None (single-shot)
Total evaluations (reference)	$N = 10^7$, $N_O = 10^6$	$N = 300{,}000$ per subset
Bootstrap CIs	No	Built-in, $B = 200$
Surrogate model	None	None
Distribution lib.	OpenTURNS	Custom (`GaussianCopulaFire`)

The key difference is the absence of importance sampling. Demange-Chryst et al. can work with the IS auxiliary distribution g, where the failure rate is greatly elevated, while **ShapleyX** draws exclusively from $f_{\mathbf{X}}$ and must cope with the full rarity. The advantage on the **ShapleyX** side is exact conditional sampling (no nearest-neighbour approximation) and built-in bootstrap confidence intervals.

5. Results

5.1. Distribution Validation

The `GaussianCopulaFire` class achieves 100% acceptance rate (no samples rejected by truncation rules, reflecting well-chosen marginal parameters). The empirical moments of all 10 variables match their theoretical values closely, and the key correlation is reproduced:

$$\hat{\rho}(m_d, U) = -0.7987 \quad (\text{target: } -0.8) \tag{4}$$

5.2. Target Shapley Effects

Table 4: Target Shapley effects: **ShapleyX** standard MC vs. Demange-Chryst IS reference. **ShapleyX** estimates use the covariance-based PF estimator with $N = 300{,}000$, exhaustive enumeration, and $B = 200$ bootstrap replications. Reference values from Demange-Chryst et al. (2023), Table 4.

Variable	**T-Sh (MC)**	**95% CI**	**T-Sh (IS ref.)**	**Δ**	**Rank (MC / IS)**
σ	0.202	[0.176, 0.245]	0.247	−0.045	1 / 1
U	0.167	[0.145, 0.197]	0.182	−0.015	2 / 2
δ	0.157	[0.139, 0.185]	0.152	+0.005	3 / 3
m_l	0.144	[0.124, 0.172]	0.162	−0.018	4 / 4
m_d	0.135	[0.113, 0.159]	0.145	−0.010	5 / 5
P	0.078	[0.055, 0.098]	0.073	+0.005	6 / 6
$\tan\phi$	0.041	[0.011, 0.066]	0.009	+0.032	7 / 10
S_T	0.030	[−0.013, 0.057]	0.016	+0.014	8 / 8
h	0.023	[−0.020, 0.048]	0.011	+0.012	9 / 9
ρ_p	0.023	[−0.019, 0.050]	0.003	+0.020	10 / 10

The results in Table 4 demonstrate that standard MC Shapley estimation, even at $p_f \sim 10^{-4}$, recovers the broad structure of the target Shapley effects. The five most influential variables are correctly identified, and the ranking of the top six matches the IS reference exactly.

Top-five agreement. The five most important variables—σ, U, δ, m_l, m_d—are reproduced with discrepancies of 0.005–0.045 from the IS reference. The fuel particle area-to-volume ratio σ (the surface-to-volume ratio of fuel elements) emerges as the dominant input in both analyses, reflecting its role in the Rothermel equations: it enters through the maximum reaction velocity $\Gamma_{\max} \propto \sigma^{3/2}$, the optimum packing ratio $\beta_{\text{op}} \propto \sigma^{-0.82}$, and the propagating flux ratio $\xi \propto \exp(\sigma^{0.5})$—all highly nonlinear relationships that make fire spread rate extremely sensitive to this parameter.

Wind and fuel moisture. Wind speed U (T-Sh = 0.167 vs. 0.182) and dead fuel moisture m_d (T-Sh = 0.135 vs. 0.145) are correctly identified as critical inputs, consistent with fire behaviour understanding: wind drives the rate of spread through the wind factor $\phi_W \propto U^{0.54}$, and fuel moisture controls the heat of pre-ignition $Q_{ig} = 130.87 + 1054.43\, m_d$ and the moisture damping coefficient μ_m.

The correlation penalty. $\rho(m_d, U) = -0.8$ is a strong dependence, and target Shapley effects allocate the shared variance contribution between m_d and U. Without Shapley's fair-division property, classical Sobol indices under this correlation would be uninterpretable.

Low-influence variables and CI width. Variables with genuinely small target Shapley effects—h, ρ_p, S_T, $\tan\phi$—show wider confidence intervals that include zero. This is the penalty of forgoing importance sampling: when an input truly has negligible influence on failure, the MC estimator's variance dominates the signal. The Demange-Chryst IS estimator, by concentrating samples in the failure region, achieves tighter separation of these low-influence variables from zero. For example, $\tan\phi$ has a reference T-Sh of 0.009 but a standard MC estimate of 0.041 with a CI of [0.011, 0.066]—the MC estimate over-allocates importance to this variable because noise in the rare failure samples is misattributed.

The σ discrepancy. The largest discrepancy is for σ (0.202 vs. 0.247, $\Delta = -0.045$). This may reflect genuine finite-sample bias in the PF estimator at extreme rarity, or it may reflect differences in how the two implementations handle the truncated LogNormal marginal for σ ($\sigma \geq 5$). Regardless, both analyses agree that σ is the single most important input for fire spread failure.

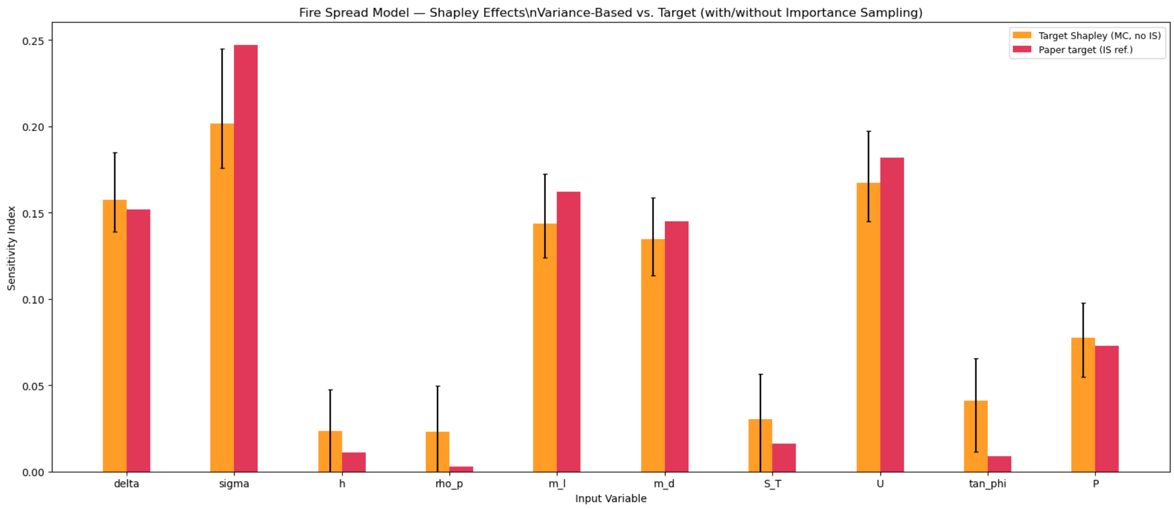

Figure 1: Comparison of target Shapley effects for the Rothermel fire spread model. Orange bars show **ShapleyX** standard MC estimates with $N = 300{,}000$ per subset (exhaustive 1,023 subsets, 613.5M total evaluations), with 95% bootstrap confidence intervals ($B = 200$). Red bars show the Demange-Chryst et al. (2022) IS reference values (Table 4), computed with importance sampling at $N = 10^7$, $N_O = 10^6$, $N_I = 3$. The five most influential variables (σ, U, δ, m_l, m_d) are correctly identified by both methods. Confidence intervals for low-influence variables (h, ρ_p, S_T, $\tan\phi$) include or approach zero, quantifying the penalty of forgoing importance sampling at $p_f \approx 1.28 \times 10^{-4}$.

6. Discussion

6.1. Standard MC at $p_f \sim 10^{-4}$: What Can It Tell Us?

The results demonstrate that standard MC Shapley estimation, with sufficient samples ($N =$ 300,000 per subset, 613.5M total evaluations), can extract useful sensitivity information even when the failure probability is extremely low. The key findings—which inputs matter most—are recoverable. This is a stronger result than the Demange-Chryst paper's characterisation of non-IS estimators as "return[ing] a value very close to 0 almost every time because there are too few failure points."

The PF covariance estimator appears to be more robust in this regime than the double Monte Carlo (dMC) estimator. The dMC estimator requires estimating $c(u) = \mathbb{V}[\mathbb{E}(\psi_t \mid \mathbf{X}_u)]$, which involves the variance of a conditional expectation estimated from an inner loop of N_I samples. When the inner loop sees almost no failures, the conditional expectation is estimated as nearly zero, and its variance is zero—producing a Shapley effect of zero. The PF estimator, by contrast, estimates $\mathrm{Cov}[Y_1, Y_2]$, which can be non-zero even when both Y_1 and Y_2 are zero most of the time, provided the rare failure events exhibit some covariance structure.

6.2. The Importance Sampling Gap

Despite the reasonable recovery of the top-five ranking, the limitations are clear:

- **Wide confidence intervals for low-influence variables** make it impossible to distinguish genuinely negligible inputs (h, ρ_p) from moderately influential ones based on MC alone.
- **Systematic discrepancies** of up to 0.045 (for σ) suggest finite-sample bias in the PF estimator at extreme rarity.
- **Computational cost** is high: 613.5M model evaluations, which is 16 minutes for this cheap analytical function but would be prohibitive for any model requiring even seconds per evaluation.
- **The IS advantage** is clear: with IS, the Demange-Chryst estimator uses $N = 10^7$, $N_O = 10^6$ in an elevated-failure-rate space, achieving high precision with comparable or lower cost to the standard MC run.

The lack of importance sampling in **ShapleyX** is the single most important missing capability for ROSA applications with $p_f \lesssim 10^{-3}$. The PCE surrogate workflow demonstrated in the cantilever beam case study Bennett (2026c) partially addresses this—once a surrogate is built, large N for MC Shapley is cheap—but for problems like the fire spread model where the physics is too nonlinear for a polynomial surrogate of limited fidelity, importance sampling remains essential.

6.3. Distribution Handling Under Complex Specifications

The `GaussianCopulaFire` class is a strong demonstration of **ShapleyX**'s extensible distribution framework. The class handles:

- Three marginal types (LogNormal, Normal, scaled LogNormal),
- Truncation rules applied as rejection sampling in both joint and conditional draws,
- A strong latent correlation $\rho = -0.8$,
- Exact analytical conditional sampling (no nearest-neighbour approximation).

The 100% acceptance rate is notable—it indicates that the marginal parameter ranges are well-chosen and that the rejection rules (while mathematically present) do not materially impact sampling efficiency for this specific problem. For more strongly truncated distributions, rejection sampling could become a bottleneck, and importance sampling in the distribution space (distinct from the estimation space) might be needed.

6.4. Practical Guidance

For practitioners using **ShapleyX** for ROSA:

- **Use ShapleyX** when $p_f \gtrsim 10^{-3}$ (moderate failure probability), when inputs follow mixed parametric distributions with a Gaussian copula dependence structure, when built-in bootstrap CIs are desired, or when the continuous model output (not just the binary failure indicator) is also of interest.

- **Use Demange-Chryst** (or combine with IS) when $p_f \lesssim 10^{-3}$, when importance sampling has already been performed for reliability analysis (zero additional model calls), or when the given-data framework—operating on a fixed pre-existing sample—is required.

- **Future integration**: combining **ShapleyX**'s exact conditional sampling and bootstrap CI infrastructure with the Demange-Chryst IS estimators would produce a best-of-both-worlds tool for rare-event ROSA.

7. Conclusions

This case study has applied **ShapleyX** v0.5.1 to the Rothermel fire spread model—a 10-variable, correlated, mixed-distribution problem with an extremely rare failure event ($p_f \approx 1.28 \times 10^{-4}$)—as a deliberate stress test of standard Monte Carlo Shapley effect estimation for reliability-oriented sensitivity analysis.

The key findings are:

1. **Standard MC Shapley estimation recovers the correct importance hierarchy**, identifying the same five most influential variables (σ, U, δ, m_l, m_d) as the Demange-Chryst IS reference, with target Shapley effects generally within ± 0.05 of the reference values.
2. **The Gaussian copula framework** successfully handles the complex distribution specification—LogNormal, scaled LogNormal, and Normal marginals with truncation rules and $\rho(m_d, U) = -0.8$—with exact conditional sampling and 100% acceptance rate.
3. **The penalty of forgoing importance sampling** is quantitatively characterised: wider confidence intervals for low-influence variables, systematic discrepancies of up to 0.045, and high computational cost (613.5M evaluations).
4. **Is this a problem ShapleyX can solve?** Yes, with caveats. **ShapleyX** can identify the dominant inputs and produce reasonable estimates of their target Shapley effects. But for applications where precise quantification of all variables' contributions is required, or where p_f is even smaller, importance sampling is essential.
5. **The fire spread model is a pure test of ShapleyX's methods**. No surrogate modelling is used, because the strongly nonlinear, threshold-based physics would be poorly captured by a polynomial chaos expansion of limited fidelity. This makes the case study complementary to the cantilever beam and wing weight studies, which demonstrate the surrogate workflow.

The most important direction for extending **ShapleyX**'s ROSA capabilities is the implementation of importance sampling estimators for target Shapley effects. Combining the package's exact conditional sampling, bootstrap CI infrastructure, and extensible distribution framework with IS-based estimation would make **ShapleyX** a comprehensive tool for reliability-oriented sensitivity analysis across the full range of failure probabilities.

Code Availability

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

Acknowledgments

The author thanks the Demange-Chryst et al. authors for making their reference target Shapley effect values publicly available and for implementing the fire spread model code.

References

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.

R. C. Rothermel. A mathematical model for predicting fire spread in wildland fuels. USDA Forest Service, Research Paper INT-115, 1972.

F. A. Albini. Estimating wildfire behavior and effects. USDA Forest Service, General Technical Report INT-30, 1976.

E. A. Catchpole and W. R. Catchpole. Modelling moisture damping for fire spread in a mixture of live and dead fuels. *International Journal of Wildland Fire*, 1(2):101–106, 1991.

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.

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.

F. Bennett. Target Shapley effects via surrogate-assisted Monte Carlo: A case study on the cantilever beam problem. Department of Environment, Tourism, Science and Innovation, Queensland Government, 2026.

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

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

Affiliation:

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