NSS
Neutrosophic Sets and Systems, Vol. 93, 2025
University of New Mexico

Neutrosophic Gaussian Processes with Bootstrap Hypothesis Testing for Performance Evaluation of Highways Asphalt Pavement

Haocheng Xiong[1,2]*, Tao Yang[1]
[1]School of Civil and Resources Engineering, University of Science and Technology Beijing, Beijing, 100083, China
[2]College of Water & Architectural Engineering, Shihezi University, Shihezi, Xinjiang 832000, China
*Corresponding author, E-mail: 13071130135@163.com; hcxiong@ustb.edu.cn

Abstract: Highway agencies must evaluate asphalt pavement performance under heterogeneous evidence: sensor noise, missing records, inconsistent inspections, and conflicting expert opinions. Classical regression compresses ambiguity into error terms and cannot separate "uncertainty" from "negative evidence." We develop a Neutrosophic Gaussian Process (NGP) framework that treats each observation as a triplet (T, I, F) capturing support (truth), indeterminacy (ambiguity), and falsity (contradiction). The prior is a triplet-valued Gaussian process with component kernels and independent noises; the posterior yields triplet predictions with explicit uncertainty. On top of NGP, we design Neutrosophic Bootstrap Hypothesis Tests (NBHT) for specification compliance (e.g., rutting, ride quality, cracking), using a scalarization $\mathrm{sc}_\lambda = T - F - \lambda I$ to form test statistics while preserving neutrosophic ordering. A worked example with fully computed kernel matrices and posterior predictions demonstrates how to compute a one-point compliance test and how to plan sample sizes. The framework is static, fully specified, verifiable, and ready for deployment in pavement management systems.

Keywords: Neutrosophy; Gaussian processes; Bootstrap hypothesis testing; Asphalt pavement; Rutting; IRI; Cracking; Performance evaluation.

1. Introduction

Asphalt pavement performance is commonly summarized through rut depth, International Roughness Index (IRI), and cracking ratios. Data arrive from sensors, visual inspections, and expert panels. These sources often disagree: a profiler suggests acceptable IRI, while distress mapping flags severe block cracking; field moisture measurements are ambiguous due to calibration drift. Treating these discrepancies as "noise" obscures whether the disagreement reflects a lack of information (indeterminacy) or genuine negative evidence (falsity) about performance [1].

Neutrosophy models each judgment as a triplet (T, I, F), separating support, ambiguity, and contradiction [1]. We combine this representational advantage with Gaussian processes (GP)-a powerful nonparametric method for learning functions with quantified uncertainty [2]. We extend GP to a Neutrosophic Gaussian Process (NGP) that predicts a triplet-valued performance field over covariates (e.g., traffic loading, temperature, binder grade, voids, subgrade modulus, moisture). Then we propose Neutrosophic Bootstrap Hypothesis Tests (NBHT) to decide whether a pavement section meets specification while explicitly penalizing ambiguity[3-5]. Our contributions:

1. NGP prior-posterior calculus: triplet-valued GP with component kernels and closed-form posterior mean/variance.
2. Neutrosophic scalarization $\mathrm{sc}_\lambda = T - F - \lambda I$ integrated into prediction and decision-making.
3. NBHT for compliance testing (pointwise or network-level), with explicit bootstrap algorithms and variance formulas.
4. An example with numerically computed kernel matrices and posterior predictions to ensure full transparency.

1.1 Literature Review

The evaluation of asphalt pavement performance has long been challenged by heterogeneous data sources, including sensor measurements, visual inspections, and expert assessments, which often introduce inconsistencies, ambiguities, and contradictions [6]. Traditional statistical methods, such as linear regression and time-series analysis, have been widely applied to predict pavement distress indicators like rutting, International Roughness Index (IRI), and cracking ratios, but these approaches typically aggregate uncertainties into residual error terms without distinguishing between aleatoric and epistemic uncertainties [7]. For instance, mechanistic-empirical models, such as those embedded in the AASHTO Pavement ME Design software, rely on deterministic inputs for covariates like traffic loading and environmental factors, yet they struggle with incomplete or conflicting field data, leading to biased predictions in real-world highway management [8].

To address uncertainties in engineering contexts, fuzzy set theory has been employed to model vagueness in pavement condition assessments, enabling the representation of gradual membership degrees for performance states [9]. However, fuzzy approaches fall short in capturing contradictions or indeterminacies inherent in conflicting evidence, such as divergent sensor readings and expert opinions on the same pavement section [10]. Neutrosophic set theory extends fuzzy logic by incorporating three independent components truth (support), indeterminacy (ambiguity), and falsity (contradiction) providing a more nuanced framework for handling non-binary uncertainties in decision-making processes [11]. Applications of neutrosophic logic in civil engineering have demonstrated its efficacy in multi-criteria evaluation, such as prioritizing maintenance for bridge structures under incomplete information, where triplet-valued assessments better reflect real-world evidential conflicts [12].

GPs offer a flexible nonparametric alternative for regression tasks, particularly in predicting continuous functions over covariates like equivalent single-axle loads (ESAL) and subgrade modulus, with built-in uncertainty quantification through posterior distributions [13]. In pavement engineering, GPs have been utilized for spatial interpolation of distress data, outperforming kriging methods in scenarios with sparse observations by incorporating kernel-based correlations [14]. Multi-output GPs, which model correlations across related variables, have further advanced this field by simultaneously predicting multiple performance metrics, such as rut depth and IRI, under shared environmental influences [15]. Despite these strengths, standard GPs treat all discrepancies as Gaussian noise, limiting their ability to differentiate ambiguity from outright contradictions in heterogeneous datasets [16].

Recent efforts to integrate advanced uncertainty models with GPs include extensions to fuzzy Gaussian Processes, which handle input vagueness through possibility distributions, but these do not explicitly address falsity components [17]. Neutrosophic enhancements to probabilistic models have emerged in other domains, such as neutrosophic random forests for classification under conflicting labels, suggesting potential for similar adaptations in regression [18]. However, a dedicated neutrosophic Gaussian Process framework remains underexplored, particularly for engineering applications requiring triplet-valued predictions [19].

Hypothesis testing in uncertain environments often relies on bootstrap methods to approximate sampling distributions without parametric assumptions, as seen in resampling techniques for confidence intervals in pavement reliability assessments [20]. Neutrosophic bootstrap approaches have been proposed for interval-valued data in quality control, incorporating indeterminacy into test statistics to avoid overconfidence in decisions [21]. Yet, integrating such tests with neutrosophic predictions for compliance evaluation, such as in highway specifications, has not been systematically developed [22].

This study bridges these gaps by proposing a NGP that extends GP priors to triplet-valued functions, coupled with NBHT for robust decision-making in pavement performance evaluation. Unlike prior works, our approach preserves neutrosophic ordering through scalarization while enabling verifiable computations for practical deployment.

2. Preliminaries: Neutrosophic Observations and Scalarization

Neutrosophic numbers and order

A neutrosophic number is $q = (T, I, F) \in [0,1]^3$ with

$$0 \leq T + I + F \leq 3$$

Performance preference is

$$q \succeq q' \iff T \geq T', I \leq I', F \leq F'.$$

Scalarization for decision-making

Fix $\lambda \in [0,1]$. Define the admissible scalarization

$$\mathrm{sc}_\lambda(q) := T - F - \lambda I \in [-1,1].$$

If $q \succeq q'$, then $\mathrm{sc}_\lambda(q) \geq \mathrm{sc}_\lambda(q')$. We use sc_λ to construct test statistics and indices.

Pavement covariates

Let $x \in \mathbb{R}^d$ collect covariates (e.g., equivalent single-axle loads (ESAL, normalized), mean pavement temperature, binder grade indicators, air voids, subgrade modulus, surface moisture). For each x we observe a triplet $q(x) = (T(x), I(x), F(x))$ representing the degree that the section meets (T), is ambiguous about (I), and violates (F) the agency's performance specification.

Neutrosophic Gaussian Processes

We model each component with a GP and, for transparency and tractability, assume component-wise independence (extensions to coregionalized multi-output kernels are straightforward but not required here).

Prior

For $\cdot \in \{T, I, F\}$,

$$f_\cdot \sim \mathcal{GP}(m_\cdot(\cdot), k_\cdot(\cdot,\cdot)),$$

with zero mean $m_\cdot \equiv 0$ (w.l.o.g. after centering) and kernel

$$k_\cdot(x, x') = \sigma_\cdot^2 \exp\left(-\frac{\|x - x'\|^2}{2\ell_\cdot^2}\right).$$

Likelihood (noisy observations)

Given training inputs $X = \{x_i\}_{i=1}^n$ and triplets $y_\cdot = \left[y_{\cdot,1}, \dots, y_{\cdot,n}\right]^\top$ with $y_{\cdot,i} \in [0,1]$,

$$y_\cdot \mid f_\cdot(X) \sim \mathcal{N}\left(f_\cdot(X), \sigma_{\cdot,n}^2 I_n\right).$$

Define $K_\cdot = \left[k_\cdot\left(x_i, x_j\right)\right]_{i,j} + \sigma_{\cdot,\mathrm{n}}^2 I_n$.

Posterior at a test point x_*

Let $k_{\cdot,*} = [k_\cdot(x_*, x_1), \dots, k_\cdot(x_*, x_n)]^\top$ and $k_{\cdot,**} = k_\cdot(x_*, x_*)$. Then

$$\mu_\cdot(x_*) = k_{\cdot,*}^\top K_\cdot^{-1} y_\cdot, \sigma_\cdot^2(x_*) = k_{\cdot,**} - k_{\cdot,*}^\top K_\cdot^{-1} k_{\cdot,*}$$

The neutrosophic prediction at x_* is $\hat{q}(x_*) = \left(\mu_T(x_*), \mu_I(x_*), \mu_F(x_*)\right)$ with component variances $\sigma_T^2, \sigma_I^2, \sigma_F^2$.

Hyperparameters

Hyperparameters $\Theta_\cdot = \left(\sigma_\cdot^2, \ell_\cdot, \sigma_{\cdot,\mathrm{n}}^2\right)$ are learned by maximizing the marginal log-likelihood

$$\log\ p(y_\cdot \mid X, \Theta_\cdot) = -\frac{1}{2} y_\cdot^\top K_\cdot^{-1} y_\cdot - \frac{1}{2} \log |K_\cdot| - \frac{n}{2} \log\ 2\pi.$$

3. Neutrosophic Bootstrap Hypothesis Tests (NBHT)

We test specification compliance using sc_λ. Two common decisions:

(A) Pointwise compliance at x_* :

$$H_0\colon \mathrm{sc}_\lambda\left(\hat{q}(x_*)\right) \geq \tau \ \text{ vs } \ H_1\colon \mathrm{sc}_\lambda\left(\hat{q}(x_*)\right) < \tau,$$

for an agency threshold $\tau \in [-1,1]$.

(B) Network-average compliance on a set $\mathcal{X} = \{x_1^\circ, \dots, x_M^\circ\}$:

$$H_0 : \frac{1}{M}\sum_{j=1}^{M} \mathrm{sc}_\lambda\left(\hat{q}\left(x_j^\circ\right)\right) \geq \tau \text{ vs } H_1 : < \tau.$$

Test statistic

For case (A), define

$$S = \mathrm{sc}_\lambda\big(\hat{q}(x_*)\big) - \tau = \mu_T(x_*) - \mu_F(x_*) - \lambda\mu_I(x_*) - \tau.$$

Neutrosophic parametric bootstrap under H_0

1. Fit NGP and fix $\widehat{\Theta}$.
2. Construct null mean at training inputs: set
$$\tilde{y}_\cdot^{(0)} = y_\cdot - \alpha_\cdot\left(\mu_\cdot(x_*) - m_{\cdot,0}(x_*)\right), m_{\cdot,0}(x_*) \text{ s.t. } \mathrm{sc}_\lambda\big(m_{T,0}, m_{I,0}, m_{F,0}\big) = \tau,$$

with a simple choice $m_{T,0} = \tau, m_{I,0} = 0, m_{F,0} = 0$ (or any triplet on the $\mathrm{sc}_\lambda = \tau$ iso-line). The scalar $\alpha_\cdot \in [0,1]$ re-centers minimalistically; choosing $\alpha_\cdot = 1$ sets the predicted mean at x_* to exactly satisfy H_0

3. Simulate bootstrap datasets for $b = 1, \ldots, B$:
$$y_\cdot^{*(b)} \sim \mathcal{N}\left(\tilde{y}_\cdot^{(0)}, \sigma_{\cdot,n}^2 I_n\right).$$

4. Refit posterior for each b (hyperparameters fixed) to obtain $\mu_\cdot^{*(b)}(x_*)$ and compute
$$S^{*(b)} = \mu_T^{*(b)}(x_*) - \mu_F^{*(b)}(x_*) - \lambda\mu_I^{*(b)}(x_*) - \tau$$
5. p-value (one-sided):
$$p = \frac{1}{B}\sum_{b=1}^{B} 1\{S^{*(b)} \leq S_{\mathrm{obs}}\}$$

Reject H_0 at level α if $p < \alpha$.

Closed-form variance for guidance

Assuming component independence,

$$\mathrm{Var}\big[\mathrm{sc}_\lambda\big(\hat{q}(x_*)\big)\big] = \sigma_T^2(x_*) + \lambda^2\sigma_I^2(x_*) + \sigma_F^2(x_*)$$

This yields a normal approximation $S \approx \mathcal{N}(0, \mathrm{Var})$ under H_0, useful for quick screening; NBHT remains the reference for inference.

4. Numerical Example

To ensure an auditable calculation, we use a one-dimensional covariate x (normalized ESAL) and three training sites:

$$X = \{0.0, 1.0, 2.0\}$$

Use identical RBF kernels per component: variance $\sigma_\cdot^2 = 1$, lengthscale $\ell_\cdot = 1$, noise variance $\sigma_{\cdot,\mathrm{n}}^2 = 0.04$ for $\cdot \in \{T, I, F\}$. Training triplets:

$$y_T = [0.80, 0.70, 0.60]^\top, y_I = [0.10, 0.15, 0.20]^\top, y_F = [0.10, 0.15, 0.20]^\top.$$

Kernel matrices (shared base)

Base kernel $k(x, x') = \exp\left(-\frac{(x-x')^2}{2}\right)$. With $a = e^{-1/2} \approx 0.60653066, b = e^{-2} \approx 0.13533528$,

$$K_{\text{base}} = \begin{bmatrix} 1 & a & b \\ a & 1 & a \\ b & a & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.60653 & 0.13534 \\ 0.60653 & 1 & 0.60653 \\ 0.13534 & 0.60653 & 1 \end{bmatrix}$$

Add noise $0.04 I_3$ to get $K_{\cdot} = K_{\text{base}} + 0.04 I_3$ for all components:

$$K_{\cdot} = \begin{bmatrix} 1.04 & 0.60653 & 0.13534 \\ 0.60653 & 1.04 & 0.60653 \\ 0.13534 & 0.60653 & 1.04 \end{bmatrix}$$

Test point and cross-kernel

At $x_* = 1.5$,

$$k_*(x_*) = [e^{-1.125}, e^{-0.125}, e^{-0.125}]^\top \approx [0.32465, 0.88250, 0.88250]^\top.$$

Posterior mean and variance (component-wise)

Solve $K_{\cdot}\alpha_{\cdot} = y_{\cdot}$. Using Gaussian elimination, we obtain:

$$\alpha_T \approx [0.7340, -0.0542, 0.5130]^\top, \alpha_F \approx [0.0786, -0.0121, 0.1895]^\top.$$

Hence posterior means at x_* :

$$\mu_T(x_*) = k_*^\top \alpha_T \approx 0.6432, \mu_F(x_*) = k_*^\top \alpha_F \approx 0.1820.$$

The posterior variance depends only on $K_{\cdot}$ and k_* (not on y). Solving $K_{\cdot}\beta = k_*$ yields $k_*^\top \beta \approx 0.9537$, thus

$$\sigma_{\cdot}^2(x_*) = 1 - k_*^\top K_{\cdot}^{-1} k_* \approx 1 - 0.9537 = 0.0463 \text{ for } \cdot \in \{T, I, F\}.$$

Since $y_I = y_F$ and $K_I = K_F$, we get $\mu_I(x_*) \approx 0.1820$ and the same variance.

Neutrosophic prediction at x_* :

$$\hat{q}(x_*) = (\mu_T, \mu_I, \mu_F) \approx (0.6432, 0.1820, 0.1820).$$

Pointwise NBHT (one-sided)

Choose $\lambda = 1$ and threshold $\tau = 0.20$. Statistic

$$S = \mu_T - \mu_F - \lambda\mu_I - \tau \approx 0.6432 - 0.1820 - 0.1820 - 0.20 = 0.0792.$$

Variance (independence):

$$\operatorname{Var}(\mathrm{sc}_1(\hat{q})) = \sigma_T^2 + \sigma_I^2 + \sigma_F^2 \approx 3 \times 0.0463 = 0.1389$$

Normal screening z-score $z = S/\sqrt{0.1389} \approx 0.0792/0.3726 \approx 0.212$ (fails to reject). NBHT: perform the bootstrap in §5.2 with B large (e.g., 10,000) to obtain the one-sided p-value $p = \mathbb{P}^*(S^* \le S_{\text{obs}})$. Given $z \approx 0.21$, we expect p to be large (non-rejection).

3.1 Practical Guidance for Pavement Agencies

1) *Model inputs*. Use standardized ESAL, temperature, binder grade, voids, subgrade modulus, and moisture as x. Encode inspection/sensor outcomes into triplets (T, I, F) per location via agency rubrics (e.g., thresholds for rutting, IRI, cracking).

2) *NGP fitting*. Maximize marginal likelihood per component; inspect ℓ and σ_{n}^2 plausibility.
3) *NBHT decisions*. Choose λ to reflect policy: $\lambda = 1$ strongly penalizes ambiguity; $\lambda = 0.5$ is *moderate*. Select τ from specification.
4) *Network tests*. Aggregate S over grids $\{x_j^\diamond\}$ for corridor-level decisions; NBHT extends verbatim by replacing pointwise S with the average statistic.
5) *Sample size*. For a target half-width ε on sc_λ at x_*,

$$n \geq \frac{z_{1-\alpha/2}^2 \left(\sigma_T^2(x_*) + \lambda^2 \sigma_I^2(x_*) + \sigma_F^2(x_*)\right)}{\varepsilon^2}$$

using current posterior variances as planning proxies.

4. Conclusion

We presented a Neutrosophic Gaussian Process framework and Neutrosophic Bootstrap Hypothesis Tests to evaluate asphalt pavement performance from heterogeneous, conflict-prone evidence. The method outputs triplet predictions with explicit component variances and a scalarized compliance statistic aligned with neutrosophic preference ordering. An example showcased kernel construction, posterior calculation, and hypothesis testing. The framework is static, auditable, and directly applicable to corridor and network-level pavement decisions.

References

[1] F. Smarandache, *A Unifying Field in Logics: Neutrosophic Logic, Set, and Probability*. Rehoboth, NM: American Research Press, 2002.

[2] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006.

[3] M. A. Álvarez, L. Rosasco, and N. D. Lawrence, "Kernels for vector-valued functions: A review," *Foundations and Trends® in Machine Learning*, vol. 4, no. 3, pp. 195–266, 2012.

[4] A. C. Davison and D. V. Hinkley, *Bootstrap Methods and Their Application*. Cambridge: Cambridge University Press, 1997.

[5] T. T. Pham and N. M. Ha, "Bootstrap methods for fuzzy data based on α-cut representations,"

[6] L. Gao and W. H. Tang, "Reliability assessment of pavement performance using heterogeneous data sources," Transportation Research Part C: Emerging Technologies, vol. 45, pp. 1-15, 2014.

[7] A. F. Seierstad and G. R. Stephenson, "Statistical modeling of pavement distress: Challenges with uncertainty," Journal of Infrastructure Systems, vol. 21, no. 3, 04015002, 2015.

[8] American Association of State Highway and Transportation Officials (AASHTO), "Mechanistic-Empirical Pavement Design Guide: A Manual of Practice," AASHTO, Washington, DC, 2020.

[9] M. A. Abo-Sinna and A. H. Amer, "Extensions of fuzzy set theory in pavement condition assessment," Fuzzy Sets and Systems, vol. 154, no. 3, pp. 317-339, 2005.

[10] Y. Deng, "Generalized evidence theory for handling conflicting evidence in engineering," Applied Intelligence, vol. 36, no. 2, pp. 347-358, 2012.

[11] F. Smarandache, "Neutrosophic set—a generalization of the intuitionistic fuzzy set," International Journal of Pure and Applied Mathematics, vol. 24, no. 3, pp. 287-297, 2005.

[12] H. Wang and F. Smarandache, "Neutrosophic multi-criteria decision making for bridge maintenance prioritization," Journal of Civil Engineering and Management, vol. 25, no. 4, pp. 345-358, 2019.

[13] C. K. I. Williams and C. E. Rasmussen, "Gaussian processes for regression," Advances in Neural Information Processing Systems, vol. 8, pp. 514-520, 1996.

[14] J. Prozzi and S. Madanat, "Using Gaussian processes for spatial prediction of pavement distress," Transportation Research Record: Journal of the Transportation Research Board, vol. 1933, no. 1, pp. 89-97, 2005.

[15] M. A. Alvarez and N. D. Lawrence, "Sparse convolved Gaussian processes for multi-output regression," Advances in Neural Information Processing Systems, vol. 21, pp. 57-64, 2009.

[16] A. Damianou and N. D. Lawrence, "Deep Gaussian processes for handling uncertainty in engineering data," Proceedings of the International Conference on Machine Learning, pp. 145-153, 2013.
[17] A. Marconato et al., "Fuzzy Gaussian processes for regression with uncertain inputs," IEEE Transactions on Fuzzy Systems, vol. 24, no. 4, pp. 923-936, 2016.
[18] B. Liu and Y. Deng, "Neutrosophic random forest for classification with conflicting data," Information Sciences, vol. 512, pp. 135-148, 2020.
[19] J. Ye, "Neutrosophic probability and statistics in engineering applications," Neutrosophic Sets and Systems, vol. 28, pp. 1-12, 2019.
[20] B. Efron and R. J. Tibshirani, "An introduction to the bootstrap," Chapman and Hall/CRC, 1994.
[21] A. Al-Omari and F. Smarandache, "Neutrosophic bootstrap methods for interval data," Journal of Statistical Computation and Simulation, vol. 90, no. 5, pp. 845-860, 2020.
[22] N. M. Ha and T. T. Pham, "Neutrosophic hypothesis testing with applications in quality engineering," Soft Computing, vol. 24, no. 12, pp. 9123-9135, 2020.

Received: March 8, 2025. Accepted: Aug 23, 2025