mathematics

MDPI

Article

Functional Regression with Reconstructed Curves: Local Rank Stability and Optimized Confidence Certificates

Wahiba Bouabsa [1,*] and Fatimah Alshahrani [2]

[1] Laboratory of Statistics and Stochastic Processes, Djillali Liabès University, Sidi Bel Abbès 22000, Algeria
[2] Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
* Correspondence: wahiba_bouab@yahoo.fr

Abstract

Reconstructing functional predictors can change nearest-neighbor ranks and introduce regression bias. Our contribution is a set of deterministic and probabilistic results quantifying these effects for an existing rank-weighted smoother. We distinguish direct pointwise inference, same-sample equivalence to the exact-curve estimator, and deterministic prediction-perturbation certification. A joint assignment relaxation bounds fitted-value changes. Under local radial-density and polynomial small-ball assumptions, local crowding gives a sufficient rank-stability rate and resolution condition; this is not a general infinite-dimensional rate. We adapt established threshold weighting to certified intervals with finite-sample coverage only when smoothness, conditional noise, and reconstruction-error upper bounds are valid independently of the response errors used in the interval. The numerical study comprises 20,175 simulated datasets, including irregular grids, Laplace measurement errors, and population principal-component truncation, with paired Monte Carlo uncertainty and neighborhood sensitivity analyses. Certified intervals remain substantially conservative, including after radius optimization. Predictive rankings vary with target location, and PCR outperforms the tested neighbor methods on Tecator. The Tecator experiment audits controlled coarsening of measured spectra; it does not validate inference for an unobserved continuous trajectory or provide prospective reconstruction calibration.

Keywords: functional regression; nearest neighbors; reconstructed curves; rank perturbation; small-ball probability; confidence intervals; spectroscopy

MSC: 62G08; 62R10; 62G15

Check for updates

Academic Editor: Bo Wang

Received: 20 September 2026
Revised: 20 October 2026
Accepted: 26 October 2026
Published: 30 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Scalar-on-function regression relates a real response to a curve or other functional predictor. Nonparametric local methods are attractive because they avoid a finite-dimensional linear regression equation, but their behavior is driven by the geometry of small neighborhoods. In applications, those neighborhoods are usually computed from reconstructed curves rather than latent trajectories. Small reconstruction errors can therefore affect both the radius of a neighborhood and the identities or ranks of the nearest observations. The main contribution is not a new estimator. It is a collection of deterministic and probabilistic results quantifying reconstruction effects on an existing rank-weighted functional regression rule, together with prediction certificates conditional on a valid reconstruction bound and statistical intervals requiring externally valid smoothness, noise, and reconstruction bounds.

Functional nearest-neighbor and kernel regression have substantial existing literature. Laloë [1], Burba et al. [2], Biau et al. [3], Kara-Zaitri et al. [4], and Kudraszow and Vieu [5] establish consistency, rates, or data-driven neighborhood selection for functional kNN procedures. Ferraty and Vieu [6] and Ferraty et al. [7] develop the broader nonparametric functional framework. Earlier work on nearest-neighbor weighting also includes Carroll and Härdle [8]. These results make clear that rates depend on functional concentration. Under polynomial small-ball behavior, one obtains power rates, whereas genuinely infinite-dimensional designs can have much thinner small-ball probabilities. Mas [9] derives lower bounds tied directly to small-ball probabilities, and Meister [10] obtains logarithmic minimax rates under metric-entropy conditions. Accordingly, the polynomial regime used below is an explicit local effective-dimension assumption, not a generic description of all functional data.

Reconstruction and discretization have also been studied previously. Forzani et al. [11] show how consistency and rates can be transferred from fully observed trajectories to discretely sampled or reconstructed curves. The present paper addresses a different question: conditional on a realized reconstruction-error bound, how much can the nearest-neighbor ordering and the fitted prediction change on the same sample? Xing, Song, and Cheng [12] study classical kNN under random input perturbation and characterize asymptotic predictive regret and a perturbation phase transition. Their setting is complementary to ours: it does not provide observation-specific deterministic rank intervals or a realized-response assignment bound for reconstructed functional covariates.

Formal verification of kNN stability is also established. Fassina, Ranzato, and Zanella [13] use abstract interpretation and discuss regression extensions in Section 3.3 of their own article (not of the present paper). Here, Theorem 2 specifies a feasible rank-permutation family from distance-error allowances and bounds the change of a rank-weighted prediction on the same realized responses. Its endpoints are exact for that relaxation, not for every realizable curve perturbation. Theorem 7 separately connects reconstruction resolution to local rank stability under stated design assumptions. We claim these specified bounds and conditions, not the general idea of certifying kNN.

The finite-sample inference component belongs to the bias-aware confidence-interval literature. Armstrong and Kolesár [14] optimize inference under smoothness restrictions, while Schennach [15] develops confidence intervals that explicitly account for non-negligible nonparametric bias rather than removing it by undersmoothing. Anava and Levy [16] already derive threshold weights by minimizing a linear bias bound plus an ℓ_2 noise penalty over the simplex. Theorem 9 adapts this established optimization to reconstructed distances, Hölder bias allowances, and conditional sub-Gaussian errors, with explicit reconstruction-failure probability.

Empirical comparison also requires care. Elmakias and Vilenchik [17] emphasize that benchmark conclusions depend strongly on dataset difficulty; this is relevant here because a single central simulation target and a single spectroscopy dataset cannot establish broad predictive superiority. We therefore add random and tail-region targets, tune local methods using training data, and include functional kernel and PCR baselines. In the Tecator benchmark, PCR is used as a predictive representation rather than interpreted as a set of physically meaningful axes. This distinction is important in light of Vilenchik, Yichye and Abutbul [18], who show that semantic interpretation of principal components can itself be unstable.

The paper separates three scientific questions that are sometimes conflated. Direct inference asks whether the reconstructed estimator itself is asymptotically valid. Paired equivalence asks whether reconstructed and exact-curve estimators are asymptotically interchangeable on the same dataset. Prediction certification asks for a deterministic bound

on the fitted-value change under an externally supplied reconstruction-error tolerance. Table 1 summarizes the targets and outputs.

Table 1. Three distinct contributions and the objects they control.

Component	Statistical Target	Main Additional Requirement	Practical Output
Direct reconstructed inference	$m(x)$ via $\widetilde{m}_k$	localization and $\delta_n \to 0$; undersmoothing only for bias omission	risk bound and studentized pointwise limit
Same-sample equivalence	$\widetilde{m}_k - \widehat{m}_k$	local rank displacement small relative to k	condition under which exact and reconstructed estimators are interchangeable
Deterministic certification	change in a fitted prediction	externally valid finite-sample distance-error bound δ	auditable upper bound; inconclusive when larger than the user's tolerance

The paper focuses on the deterministic rank-assignment construction in Theorem 2 and the local resolution condition in Theorem 7. Theorem 9 complements these results by adapting an existing weight-optimization principle to the reconstruction experiment. Table 2 distinguishes the relevant prior results from the additional conclusions proved here.

Table 2. Relationship of the principal results to prior work.

Result	Closest Existing Line	Additional Issue Addressed	Result Supplied Here
Theorem 2	kNN under perturbation [12]; reconstructed curves [11]; stability verification [13]	realized, observation-specific feasible rank intervals jointly constrained by a permutation; deterministic fitted-value extrema conditional on observed responses	two assignment problems giving exact extrema over the stated feasible rank-permutation relaxation, and a bound no larger than the separate-rank relaxation
Theorem 7	functional kNN rates [2,3]; discretization transfer [11]; small-ball/minimax rates [9,10]	a same-sample rank-displacement rate that directly links reconstruction resolution to oracle equivalence	under polynomial local small-ball density, $\rho_n^{\mathrm{loc}} = O_p(\delta_n/h_n + \log n/k)$ and an explicit sufficient resolution condition; no claim is made for non-polynomial small-ball regimes
Theorem 9	threshold weights [16], problem P2 and Theorem 3.1; bias-aware inference [14,15]	Hölder bias allowances from reconstructed distances and a calibrated reconstruction-failure probability	adaptation of the established threshold optimizer; coverage at least $1 - \alpha - \eta$ for the stated reconstructed-data experiment

Our experiments test the limits of these distinctions rather than a universal prediction claim. The original simulation has 27 cells and 10,800 generated datasets with fixed k to isolate weighting effects; additional experiments tune k, vary target location, add functional kernel regression and PCR, stress the calibration constants, and examine three alternative reconstruction mechanisms. Paired Monte Carlo standard errors use the independent training replication as their unit. The Tecator study removes duplicate spectra before splitting, tunes all predictive baselines within training folds, and reports the full distribution of certificate slack rather than averages alone.

The remainder of the paper is organized as follows. Section 2 introduces the regression model and rank weights. Section 3 develops deterministic stability and the joint assignment

certificate. Section 4 gives risk and pointwise inference. Section 5 treats paired oracle equivalence and local rank stability. Section 6 develops finite-sample certificates and calibration. Section 7 presents simulations. Section 8 reports Tecator. Sections 9 and 10 discuss limitations and conclusions.

2. Model, Reconstruction, and Rank Weights

Let $(\mathcal{H}, \|\cdot\|)$ be a separable Hilbert space, and fix a target $x \in \mathcal{H}$. We observe independent and identically distributed pairs (X_i, Y_i) satisfying

$$Y_i = m(X_i) + \varepsilon_i, \qquad \mathbb{E}(\varepsilon_i \mid X_i) = 0, \quad \mathbb{E}(\varepsilon_i^2 \mid X_i) = \sigma^2(X_i). \tag{1}$$

Write $D_i = \|X_i - x\|$. Exact distances are assumed distinct almost surely when statements about their ranks are made. Let $r_i \in \{1, \dots, n\}$ denote their ranks, and let $D_{(j)}$ be the jth order statistic. Reconstruction produces $\widetilde{X}_i, \tilde{x}$, distances $\widetilde{D}_i = \left\|\widetilde{X}_i - \tilde{x}\right\|$, and ranks $\tilde{r}_i$. Reconstructed ties are broken by the original observation index. Neither the reconstructor nor this tie rule uses responses.

For precision, let $\mathcal{G}_n$ contain all latent predictors, all predictor measurement errors, the fixed target and its measurement errors, and all reconstructions. Throughout conditional inference arguments, assume that the response errors are independent given $\mathcal{G}_n$, with conditional mean zero and conditional variance $\sigma^2(X_i)$. This holds, for example, when predictor measurement errors are independent of the response errors given the latent predictors. A response-trained representation need not satisfy this assumption.

For $2 \le k < n$, define $a_{j,k} = (1 - j/(k+1))_+$, where $u_+ = \max(u, 0)$, and

$$S_k = \sum_{j=1}^{n} a_{j,k} = \frac{k}{2}, \qquad w_i = \frac{a_{r_i,k}}{S_k}, \quad \widetilde{w}_i = \frac{a_{\tilde{r}_i,k}}{S_k}, \qquad \widehat{m}_k(x) = \sum_i w_i Y_i, \quad \widetilde{m}_k(x) = \sum_i \widetilde{w}_i Y_i. \tag{2}$$

Exactly k observations receive positive weight. The slight shift from k to $k+1$ ensures that the kth observation has positive weight, including at small k.

Lemma 1 (Normalization and empirical-mass representation)**.** *Both weight vectors in* (2) *sum to one and have maximum* $2/(k+1)$*. Their common squared norm is*

$$V_k = \sum_i w_i^2 = \sum_i \widetilde{w}_i^2 = \frac{2(2k+1)}{3k(k+1)}, \qquad kV_k \longrightarrow \frac{4}{3}. \tag{3}$$

If $\hat{F}_x(t) = n^{-1} \sum_i \mathbf{1}\{D_i \le t\}$ *and* $q = (k+1)/n$*, then* $a_{r_i,k} = K(\hat{F}_x(D_i)/q)$ *for* $K(u) = (1-u)_+$*. The estimator is invariant under any strictly increasing transformation of all distances, provided that the tie rule is unchanged.*

Proof. The positive numerators are $k/(k+1), (k-1)/(k+1), \dots, 1/(k+1)$. Their sum is $k/2$. Dividing their squared sum by $k^2/4$, and using $\sum_{j=1}^{k} j^2 = k(k+1)(2k+1)/6$, gives (3). The largest normalized numerator is $2/(k+1)$. In the absence of ties, $\hat{F}_x(D_i) = r_i/n$, proving the representation. A strictly increasing transformation preserves the complete ordering, and hence every weight. These identities do not require a density for X. □

This representation identifies the method as a rank-weighted neighbor smoother. Uniform kNN has the same monotone invariance. Invariance does not make the choice of distance irrelevant: two different metrics may induce different orderings. Nor does it remove the small-ball difficulty, because localization still depends on the distribution of D_i.

For $\delta \ge 0$ define the event

$$\mathcal{E}_\delta = \{\max_i |\widetilde{D}_i - D_i| \le \delta\}. \tag{4}$$

The triangle inequality guarantees this event whenever $\max_i \left\|\widetilde{X}_i - X_i\right\| + \|\tilde{x} - x\| \le \delta$. The target reconstruction error must be included; otherwise, an interval may target the wrong curve even when the training sample is large. Table 3 summarizes the notation used in the results below.

Table 3. Notation and the status of the principal quantities.

Symbol	Meaning
$X_i, x;\ \widetilde{X}_i, \tilde{x}$	Latent predictors and fixed latent target; their reconstructions.
$Y_i, \varepsilon_i, m(x)$	Observed response, response error, and target conditional mean.
$D_i, \widetilde{D}_i;\ r_i, \tilde{r}_i$	Exact and reconstructed target distances; their ranks.
$D_{(j)}, \widetilde{D}_{(j)}$	Distance order statistics.
k, S_k, V_k	Neighborhood size, rank-profile normalization, squared weight norm.
$\widehat{m}_k, \widetilde{m}_k$	Exact-curve and reconstructed rank-weighted predictions.
$\delta, \mathcal{E}_\delta;\ \delta_n$	Finite-sample distance-error allowance and its validity event; deterministic asymptotic allowance. The target error is included.
$\ell_i, u_i;\ q_i^{\mathrm{rank}}$	Observable feasible exact-rank limits; rank-weight perturbation allowance in (7).
$C_k, C_k^{\mathrm{joint}}$	Separate-rank and joint-assignment prediction-perturbation bounds.
F_x, H_x, h_n	Radial distribution, its inverse, localization radius $H_x(2k/n)$.
$R_n, \rho_n;\ R_n^{\mathrm{loc}}, \rho_n^{\mathrm{loc}}$	Global and local maximum rank displacements and their ratios to k.
$L, \beta, \bar{\sigma}$	Smoothness bound, Hölder exponent, variance or sub-Gaussian scale bound, as specified for each theorem.
$b_i, A(w), t$	Smoothness allowance $L(\widetilde{D}_i + \delta)^\beta$, certified half-width objective, optimal threshold.
$\alpha, \eta;\ \mathcal{G}_n$	Response-tail and reconstruction-failure budgets; conditioning information excluding current response errors.

The results should be read as a theorem-by-theorem map of reconstruction effects. Table 4 classifies their role; the rank-weighted estimator itself is an existing type of neighbor smoother, and the threshold optimization is attributed to [16]. Novelty is claimed for the stated reconstruction-specific deterministic certificates and sufficient local resolution analysis, not for generic kNN consistency, a new regression estimator, or the threshold principle.

Table 4. Theorem-by-theorem contribution map.

Result	Role	Status of Contribution
Theorem 1	Rank displacement, localization, observable perturbation bound	Reconstruction-specific deterministic formulation from elementary ordering and convex-weight inequalities.
Theorem 2	Joint fitted-value extrema	Specified feasible-rank permutation certificate; assignment optimization and general kNN certification [13] are established.
Theorem 3	Exact-curve risk	Reference result from standard local bias–variance arguments.
Theorem 4	Exact-curve studentization	Reference conditional limit argument under localization and moment conditions.
Theorem 5	Direct reconstructed inference	Reconstruction-specific localization transfer; separates direct inference from rank equivalence.
Theorem 6	Same-sample equivalence	Explicit rank-to-weight perturbation bounds and sufficient equivalence criterion for the stated taper.
Theorem 7	Local rate and resolution	Proposed sufficient local-crowding bound under polynomial small-ball and radial-density assumptions.
Theorem 8	Finite-grid certified intervals	Reconstruction-adjusted bias and tail bound; standard union bound gives selection protection.
Theorem 9	Radius-optimal certified weights	Adaptation of [16]; the contribution is the reconstruction-adjusted coverage statement, not the optimizer formula.

3. Deterministic Stability and Observable Certificates

Our first result holds for any realized responses, including unbounded response models. Its certificate is finite for each finite sample.

Theorem 1 (Rank displacement, localization, and certification)**.** *On $\mathcal{E}_\delta$, the following statements hold.*

(i) $|\widetilde{D}_{(j)} - D_{(j)}| \le \delta$ *for every* j. *If* $\tilde{r}_i \le k$, *then* $D_i \le D_{(k)} + 2\delta$ *and* $D_i \le \widetilde{D}_{(k)} + \delta$.

(ii) *Let* $M_i(\delta) = \sum_{j \ne i} \mathbf{1}\{|D_j - D_i| \le 2\delta\}$. *Then* $|\tilde{r}_i - r_i| \le M_i(\delta)$. *In particular, if* $D_{(k+1)} - D_{(k)} > 2\delta$, *the set of the first* k *neighbors is unchanged. Their internal order need not be unchanged.*

(iii) *Define observable rank limits*

$$\ell_i = 1 + \sum_{j \ne i} \mathbf{1}\{\widetilde{D}_j < \widetilde{D}_i - 2\delta\}, \tag{5}$$

$$u_i = 1 + \sum_{j \ne i} \mathbf{1}\{\widetilde{D}_j \le \widetilde{D}_i + 2\delta\}. \tag{6}$$

Then $\ell_i \le r_i \le u_i$. *Put*

$$q_i^{\mathrm{rank}} = \frac{\max\{|a_{\ell_i,k} - a_{\tilde{r}_i,k}|, |a_{u_i,k} - a_{\tilde{r}_i,k}|\}}{S_k}. \tag{7}$$

Writing $\mathcal{R}_Y = \max_i Y_i - \min_i Y_i$, *we have*

$$|\widetilde{m}_k(x) - \widehat{m}_k(x)| \le C_k(\delta, Y) := \mathcal{R}_Y \min\left\{1, \frac{1}{2}\sum_i q_i^{\mathrm{rank}}\right\}. \tag{8}$$

Proof. At least j exact distances are no larger than $D_{(j)}$, so at least j reconstructed distances are no larger than $D_{(j)} + \delta$. Thus $\widetilde{D}_{(j)} \le D_{(j)} + \delta$. Interchanging exact and reconstructed distances yields the opposite inequality. For a selected reconstructed neighbor, $D_i \le \widetilde{D}_i + \delta \le \widetilde{D}_{(k)} + \delta \le D_{(k)} + 2\delta$, proving (i).

If $D_j < D_i - 2\delta$, then $\widetilde{D}_j \le D_j + \delta < D_i - \delta \le \widetilde{D}_i$. If $D_j > D_i + 2\delta$, the reconstructed order is also fixed in the opposite direction. Hence, only the $M_i(\delta)$ close distances can move across observation i, proving the rank bound. For any exact inside neighbor and exact outside neighbor, their distance difference is at least $D_{(k+1)} - D_{(k)}$. When this exceeds 2δ, no such pair can reverse order, proving set preservation.

For (iii), if $\widetilde{D}_j < \widetilde{D}_i - 2\delta$, then $D_j \le \widetilde{D}_j + \delta < \widetilde{D}_i - \delta \le D_i$. These observations must precede i in the exact ordering. If $\widetilde{D}_j > \widetilde{D}_i + 2\delta$, then $D_j > D_i$, so only observations counted in $u_i - 1$ can precede i. This proves the limits. The sequence $a_{j,k}$ is decreasing, and therefore the maximum possible difference between its value at an integer in $[\ell_i, u_i]$ and its value at $\tilde{r}_i$ is attained at an endpoint. Consequently, $|w_i - \widetilde{w}_i| \le q_i^{\mathrm{rank}}$.

Both weight vectors sum to one. Taking $c = (\max_i Y_i + \min_i Y_i)/2$ gives

$$|\widetilde{m}_k - \widehat{m}_k| = \left|\sum_i (\widetilde{w}_i - w_i)(Y_i - c)\right| \le \frac{\mathcal{R}_Y}{2}\sum_i q_i^{\mathrm{rank}}.$$

Each estimator is also in the interval $[\min_i Y_i, \max_i Y_i]$, so their difference cannot exceed $\mathcal{R}_Y$. Taking the smaller bound proves (8). □

The certificate concerns the difference between two predictions on the same responses. It is not a confidence interval for $m(x)$. It may be conservative because the separate rank intervals do not enforce all constraints of a single joint ordering. Conversely, a small neighbor-set change does not ensure a small rank-weighted change unless internal ranks are also controlled. For uniform kNN, the boundary gap condition alone preserves the prediction.

A Joint Assignment Certificate

The intervals $[\ell_i, u_i]$ in Theorem 1 can be enforced jointly rather than bounded separately. Let $v_j = a_{j,k}/S_k$ and let Π_δ be the set of permutations π of $\{1, \ldots, n\}$ satisfying $\ell_i \le \pi(i) \le u_i$ for every i. Define

$$P_- = \min_{\pi \in \Pi_\delta} \sum_i Y_i v_{\pi(i)}, \qquad P_+ = \max_{\pi \in \Pi_\delta} \sum_i Y_i v_{\pi(i)}.$$

Theorem 2 (Joint permutation relaxation)**.** *The set Π_δ is nonempty. On $\mathcal{E}_\delta$,*

$$|\widetilde{m}_k - \widehat{m}_k| \le C_k^{\text{joint}} := \max\{P_+ - \widetilde{m}_k, \widetilde{m}_k - P_-\} \le C_k(\delta, Y). \tag{9}$$

The endpoints are solutions of two linear assignment problems with costs $Y_i v_j$ and $-Y_i v_j$, respectively, forbidding entries $j \notin [\ell_i, u_i]$.

Proof. For any $\delta \ge 0$, the reconstructed rank $\tilde{r}_i$ lies between ℓ_i and u_i, including under index-based tie breaking: the lower count includes only strictly smaller distances and the upper count includes all tied distances. Hence the reconstructed permutation is feasible and $P_- \le \widetilde{m}_k \le P_+$. On $\mathcal{E}_\delta$, the exact permutation is feasible by Theorem 1, and thus $P_- \le \widehat{m}_k \le P_+$. This proves the first inequality.

For every feasible permutation, monotonicity of the profile gives $|v_{\pi(i)} - v_{\tilde{r}_i}| \le q_i^{\text{rank}}$ with q_i^{rank} defined in (7). Both assigned weight vectors sum to one. Subtracting the midrange of the responses, as in the proof of (8), bounds each feasible prediction difference by $\mathcal{R}_Y \sum_i q_i^{\text{rank}}/2$. Each prediction is also a convex combination of the responses, so the difference is at most $\mathcal{R}_Y$. Apply these two bounds to the extremal assignments to obtain the second inequality.

A feasible permutation selects exactly one allowed column per row and one row per column. Its total assignment cost is precisely the displayed objective, which proves the computational characterization. The assignment formulation is a standard optimization problem; the certificate follows from embedding the unknown true ordering into its feasible set. □

The bound is exact for the stated permutation relaxation. Not every permitted permutation needs to arise from a common collection of distances within the error bands, so it is not claimed to be the sharp bound over all physically realizable reconstructions. The rank-weight allowance q_i^{rank} in (7) and the smoothness allowance b_i in Theorem 9 are different quantities.

Abstract-interpretation enclosures [13] motivate specifying the uncertainty set explicitly. In (9), the endpoints jointly use the realized responses over the allowed rank permutations. This does not establish sharpness over realizable curve perturbations or numerical superiority over other verification frameworks.

Implementation cost

Sorting distances and computing all rank limits by binary searches take $O(n \log n)$ time. The dense assignment implementation stores $n \times n$ costs and a feasibility mask, using $O(n^2)$ memory, and solves two square assignment problems with cubic worst-case time. It uses SciPy's modified Jonker–Volgenant solver; infeasible positions have infinite cost. Reconstructed ranks supply a feasible permutation even with ties. This cost can be appreciable for large training samples; no sparse or faster worst-case algorithm is claimed. A measured Tecator benchmark appears in Section 8.6.

4. Risk and Pointwise Inference

Conditioning, Randomness, and Scope of the Assumptions

The latent predictors $X_1, \ldots, X_n$ are random before conditioning; x is a fixed point of $\mathcal{H}$. Predictor measurement errors and the resulting reconstructions are also random. The sigma-field $\mathcal{G}_n$ contains these predictor quantities, the target measurement information, and any independent calibration information, but excludes the response errors ε_i used to form the interval. Conditional on $\mathcal{G}_n$, the latent and reconstructed distances, their ranks, the rank weights for a fixed k, and any predictor-measurable tolerance are fixed. Response

errors remain random, conditionally independent, and centered, with the stated conditional moments. The function m and the true variance function are unknown deterministic features of the model.

A reconstruction may use predictor measurements or an independently trained representation if the conditional error assumptions still hold. A reconstruction trained using the current analysis responses generally does not satisfy the required centering and independence; it is excluded from these inference statements unless a separate argument establishes those properties. Merely calling a procedure cross-fitted does not establish them. The deterministic prediction-perturbation theorems, in contrast, concern arbitrary realized responses and require no response-noise law.

For asymptotic statements, $k = k_n$ and δ_n are deterministic sequences; h_n is a deterministic design quantile. Cross-validation used in the predictive experiments does not automatically inherit these fixed-sequence limit theorems. The finite-grid simultaneous guarantee in Theorem 8 separately protects response-dependent selection of k from its stated grid. Theorem 9 instead requires weights chosen without the current response errors.

The constants L, β, and $\bar{\sigma}$ are assumptions in the asymptotic results, not estimated quantities there. To report certified intervals, they must be known structural bounds or externally calibrated upper bounds, together with a valid δ. Ordinary same-sample estimates are not automatically upper bounds; failure probabilities of independent calibration must be included in the coverage budget. Table 5 records what fails when the main assumptions do not hold.

Table 5. Assumptions, their uses, and limitations.

Assumption	Used for	Consequence of Violation
Local Hölder control and localization	Risk and bias bounds	Selected curves need not have small regression bias; bias omission can fail.
Conditional centering and independence of response errors	Risk, Gaussian limits, exponential tails	The stated variance and tail calculations no longer apply; response-informed reconstruction can introduce selection bias.
Positive continuous variance and local fourth moments	Studentization	Residual scale consistency or the conditional Gaussian limit can fail.
High-probability $\delta_n \to 0$; undersmoothing for bias omission	Direct asymptotic inference	Fixed noisy resolution does not justify a bias-free mean interval.
Polynomial small balls and local radial-density bound	Explicit local rate in Theorem 7	The displayed power resolution condition is not available in general; the deterministic certificates remain valid on their error event.
Externally valid smoothness, sub-Gaussian noise, and reconstruction bounds	Certified mean intervals	Empirical coverage or fitted constants alone do not establish the stated finite-sample coverage.

Assume that for some $r_0 > 0$, $0 < \beta \le 1$ and $L < \infty$,

$$|m(z) - m(x)| \le L\|z - x\|^{\beta}, \qquad \|z - x\| \le r_0. \tag{10}$$

Assume also $\sigma^2(z) \le \bar{\sigma}^2$ on this ball. These are local assumptions; certified finite-sample mean intervals below require valid global bounds or a certified radius within r_0.

Define $F_x(h) = \mathbb{P}(D_i \le h)$ and $H_x(q) = \inf\{h : F_x(h) \ge q\}$. Suppose $F_x(h) > 0$ for every $h > 0$ and $F_x(0) = 0$. Where quantile rates are used, F_x is continuous and strictly increasing near zero. For the asymptotic results, the neighborhood sequence is deterministic. Let $k = k_n \to \infty$, $k/n \to 0$, and put $h_n = H_x(2k/n)$ for all sufficiently large n.

Lemma 2 (Small-ball localization)**.** *Under these assumptions, $h_n \to 0$, $D_{(k)} \to 0$ in probability, and*

$$\mathbb{P}\{D_{(k)} > h_n\} \le e^{-k/4}. \tag{11}$$

Proof. For each fixed $h > 0$, eventually $2k/n < F_x(h)$, and therefore $h_n \le h$. This proves $h_n \to 0$. The count $N_n = \sum_i \mathbf{1}\{D_i \le h_n\}$ is binomial with mean $2k$. The event $D_{(k)} > h_n$ implies $N_n < k$. The multiplicative Chernoff bound at one half of the mean gives $\mathbb{P}(N_n < k) \le \exp\{-(2k)(1/2)^2/2\} = e^{-k/4}$. Finally, $D_{(k)} \le h_n$ with probability tending to one, giving localization. □

Theorem 3 (Exact-curve conditional risk and rate)**.** *On* $\{D_{(k)} \le r_0\}$,

$$\mathbb{E}[(\widehat{m}_k(x) - m(x))^2 \mid \mathcal{G}_n] \le L^2 D_{(k)}^{2\beta} + \bar{\sigma}^2 V_k. \tag{12}$$

Under Lemma 2,

$$\widehat{m}_k(x) - m(x) = O_p(h_n^\beta + k^{-1/2}). \tag{13}$$

If $F_x(h) \sim c_x h^d$ *for* $d > 0$*, this becomes* $O_p((k/n)^{\beta/d} + k^{-1/2})$*. Balancing the two terms yields* $k \asymp n^{2\beta/(2\beta+d)}$ *and rate* $n^{-\beta/(2\beta+d)}$*.*

Proof. Conditionally on $\mathcal{G}_n$, the weights and predictors are fixed. Decompose

$$\widehat{m}_k - m(x) = B_k + Z_k, \quad B_k = \sum_i w_i\{m(X_i) - m(x)\}, \quad Z_k = \sum_i w_i \varepsilon_i.$$

All positive weights correspond to distances no larger than $D_{(k)}$. Positivity and normalization imply $|B_k| \le L D_{(k)}^\beta$. Conditional independence and zero means imply $\mathbb{E}(Z_k \mid \mathcal{G}_n) = 0$ and $\mathbb{E}(Z_k^2 \mid \mathcal{G}_n) = \sum_i w_i^2 \sigma^2(X_i) \le \bar{\sigma}^2 V_k$. The cross term vanishes, proving (12).

On $D_{(k)} \le h_n$, the bias is at most $L h_n^\beta$. Conditional Chebyshev gives $\mathbb{P}(|Z_k| > A/\sqrt{k}, D_{(k)} \le h_n) \le k\bar{\sigma}^2 V_k / A^2$. Since kV_k is bounded and the excluded event has probability at most $e^{-k/4}$, the asserted stochastic order follows. Inverting $F_x(h) \sim c_x h^d$ gives $h_n \sim (2k/(nc_x))^{1/d}$. Solving $(k/n)^{\beta/d} = k^{-1/2}$ gives the displayed choice and rate. This balancing uses unknown smoothness and concentration parameters and is not an adaptation theorem. □

Polynomial small-ball behavior is only one regime. For example, if $F_x(h) \sim \exp(-ch^{-a})$ with $a, c > 0$, then $h_n \asymp \{\log(n/k)\}^{-1/a}$, and the bias order is logarithmic. Rank smoothing cannot convert such an infinite-dimensional concentration regime into a polynomial one. This distinction parallels the logarithmic lower-bound behavior in Mas [9] and the logarithmic minimax rates established by Meister [10]. Theorem 3 therefore gives a transparent polynomial-regime benchmark, not a claim that polynomial rates are generic in functional regression.

Theorem 4 (Exact-curve studentization)**.** *In addition to the local assumptions above, suppose* $\sigma^2(z) \to \sigma^2(x) > 0$ *as* $z \to x$*, and that* $\mathbb{E}(\varepsilon_i^4 \mid \mathcal{G}_n) \le C$ *whenever* $D_i \le r_0$*. Define*

$$\hat{s}_k^2 = \sum_i w_i (Y_i - \widehat{m}_k)^2. \tag{14}$$

Then $\hat{s}_k^2 \to \sigma^2(x)$ *in probability and*

$$\frac{\widehat{m}_k - m(x) - B_k}{\hat{s}_k \sqrt{V_k}} \Rightarrow N(0,1). \tag{15}$$

If additionally $\sqrt{k} h_n^\beta \to 0$*, the bias* B_k *can be omitted. In this case* $\widehat{m}_k \pm z_{1-\alpha/2} \hat{s}_k \sqrt{V_k}$ *is an asymptotic* $(1-\alpha)$ *confidence interval for the conditional mean* $m(x)$*.*

Proof. Put $v_n^2 = \sum_i w_i^2 \sigma^2(X_i)$. Since selected distances converge to zero,

$$\left| \frac{v_n^2}{V_k} - \sigma^2(x) \right| \le \sup_{\|z-x\| \le D_{(k)}} |\sigma^2(z) - \sigma^2(x)| \longrightarrow_p 0.$$

In particular, with probability tending to one, $v_n^2 \geq \sigma^2(x)V_k/2$. On that event, the conditional Lyapunov ratio satisfies

$$\frac{\sum_i w_i^4 \mathbb{E}(\varepsilon_i^4 \mid \mathcal{G}_n)}{v_n^4} \leq C' \frac{(\max_i w_i)^2 V_k}{V_k^2} = O(k^{-1}).$$

Thus the conditional characteristic function of Z_k/v_n converges in probability to $e^{-t^2/2}$ for each t. To justify this for random conditioning arrays, take an arbitrary subsequence and extract a further subsequence on which localization, the variance ratio, and the Lyapunov bound hold almost surely. The ordinary triangular-array Lyapunov theorem applies to each such conditional array. Boundedness of characteristic functions then gives unconditional convergence. Since this argument applies to every subsequence, $Z_k/v_n \Rightarrow N(0,1)$.

For variance estimation, write $b_i = m(X_i) - m(x)$ and $A_k = \widehat{m}_k - m(x)$. The exact identity

$$\hat{s}_k^2 = \sum_i w_i(\varepsilon_i + b_i)^2 - A_k^2$$

holds because $\sum_i w_i = 1$. Conditional on $\mathcal{G}_n$, the variance of $\sum_i w_i \varepsilon_i^2$ is at most $CV_k = O(k^{-1})$, and its conditional expectation $\sum_i w_i \sigma^2(X_i)$ tends to $\sigma^2(x)$. Also $\sum_i w_i b_i^2 \leq L^2 D_{(k)}^{2\beta} = o_p(1)$. The mixed term has magnitude at most $2(\sum_i w_i \varepsilon_i^2)^{1/2}(\sum_i w_i b_i^2)^{1/2} = o_p(1)$ by Cauchy–Schwarz. Theorem 3 gives $A_k = o_p(1)$. This proves consistency of $\hat{s}_k^2$, and Slutsky's theorem proves (15). Finally, $|B_k| \leq Lh_n^\beta$ except on a vanishing event, and $V_k \asymp 1/k$, so undersmoothing makes the normalized bias negligible. □

Theorem 5 (Direct inference with reconstructed curves)**.** *Let $\delta_n \to 0$ be deterministic and $\mathbb{P}(\mathcal{E}_{\delta_n}) \to 1$. On $\mathcal{E}_{\delta_n} \cap \{D_{(k)} + 2\delta_n \leq r_0\}$,*

$$\mathbb{E}[(\widetilde{m}_k - m(x))^2 \mid \mathcal{G}_n] \leq L^2(D_{(k)} + 2\delta_n)^{2\beta} + \bar{\sigma}^2 V_k. \tag{16}$$

Consequently, $\widetilde{m}_k - m(x) = O_p((h_n + \delta_n)^\beta + k^{-1/2})$. Under the moment and variance-continuity assumptions of Theorem 4, define $\widetilde{B}_k = \sum_i \widetilde{w}_i\{m(X_i) - m(x)\}$ and $\tilde{s}_k^2 = \sum_i \widetilde{w}_i(Y_i - \widetilde{m}_k)^2$. Then $\tilde{s}_k^2 \to \sigma^2(x)$ and

$$\frac{\widetilde{m}_k - m(x) - \widetilde{B}_k}{\tilde{s}_k\sqrt{V_k}} \Rightarrow N(0,1). \tag{17}$$

The bias can be omitted if $\sqrt{k}(h_n + \delta_n)^\beta \to 0$. No rank-displacement condition is required for this direct limit.

Proof. Theorem 1(i) bounds every selected exact distance by $D_{(k)} + 2\delta_n$. The weights are measurable with respect to $\mathcal{G}_n$, remain nonnegative, and have the same deterministic normalization and squared norm as the exact weights. Conditional on $\mathcal{G}_n$, the bias is bounded by $L(D_{(k)} + 2\delta_n)^\beta$ and the error variance by $\bar{\sigma}^2 V_k$. This proves (16). Conditional Chebyshev on $\mathcal{E}_{\delta_n} \cap \{D_{(k)} \leq h_n\}$ gives the stochastic rate; the complement has probability tending to zero.

For the limit distribution, the selected exact radius is $o_p(1)$. Therefore $\sum_i \widetilde{w}_i^2 \sigma^2(X_i)/V_k \to \sigma^2(x)$. The conditional Lyapunov bound is again $C'(\max_i \widetilde{w}_i)^2/V_k = O(k^{-1})$. Repeating the conditional characteristic-function argument gives the Gaussian limit for the centered error. The variance-estimation identity and its four bounds in the preceding proof remain valid after replacing $w_i, D_{(k)}, A_k$ by $\widetilde{w}_i, D_{(k)} + 2\delta_n, \widetilde{m}_k - m(x)$. They establish the consistency of $\tilde{s}_k^2$. Finally, the stated undersmoothing condition bounds $\sqrt{k}|\widetilde{B}_k|$ by a sequence tending to zero outside a vanishing event. □

The condition on δ_n is a condition on the reconstruction experiment, not a consequence of increasing n alone. For fixed noisy measurement resolution, the maximum reconstruction error typically does not vanish. Neither the bias-free interval in Theorem 5 nor same-sample oracle equivalence should then be asserted.

5. Paired Oracle Equivalence

Here "oracle" means the exact-curve estimator (2), not an estimator that knows m. Let $R_n = \max_i |r_i - \tilde{r}_i|$ and $\rho_n = R_n/k$.

Theorem 6 (Same-sample equivalence under rank stability)**.** *On* $\{R_n \le k\}$,

$$\sum_i |\widetilde{w}_i - w_i| \le 4\rho_n, \qquad \sum_i (\widetilde{w}_i - w_i)^2 \le \frac{8\rho_n^2}{k}. \tag{18}$$

If $\mathbb{P}(\mathcal{E}_{\delta_n}) \to 1$, $\rho_n \to_p 0$, *and* $h_n + \delta_n \to 0$, *then*

$$\widetilde{m}_k - \widehat{m}_k = O_p\big(\rho_n\{(h_n + \delta_n)^\beta + k^{-1/2}\}\big). \tag{19}$$

Here the notation means that the quotient by the random nonnegative expression is tight on $\rho_n > 0$, *with difference zero when* $\rho_n = 0$. *In particular, if* $\rho_n\{1 + \sqrt{k}(h_n + \delta_n)^\beta\} \to_p 0$, *then* $\sqrt{k}(\widetilde{m}_k - \widehat{m}_k) \to_p 0$.

Proof. The map $j \mapsto a_{j,k}$ is $1/(k+1)$-Lipschitz. Hence $|\widetilde{w}_i - w_i| \le 2R_n/[k(k+1)]$. A nonzero weight difference requires $r_i \le k$ or $\tilde{r}_i \le k$. In the latter case, $r_i \le k + R_n$, so there are at most $k + R_n \le 2k$ potentially nonzero differences. Multiplying the individual bound, and then its square, by $2k$ proves (18).

Write $\Delta_i = \widetilde{w}_i - w_i$. Since $\sum_i \Delta_i = 0$,

$$\widetilde{m}_k - \widehat{m}_k = \sum_i \Delta_i\{m(X_i) - m(x)\} + \sum_i \Delta_i \varepsilon_i.$$

On $\mathcal{E}_{\delta_n}$, every index in the union of the two supports lies at an exact distance of at most $D_{(k)} + 2\delta_n$. The deterministic conditional mean difference is thus at most $4L\rho_n(D_{(k)} + 2\delta_n)^\beta$. The conditional variance of the second sum is at most $8\bar{\sigma}^2\rho_n^2/k$. For $\rho_n > 0$, conditional Chebyshev gives

$$\mathbb{P}\left(\left|\sum_i \Delta_i \varepsilon_i\right| > A\rho_n/\sqrt{k} \mid \mathcal{G}_n\right) \le 8\bar{\sigma}^2/A^2$$

on the localization event. When $\rho_n = 0$, the weights agree, and both terms are zero. Lemma 2 and the assumed high-probability events prove (19). Multiplication by $\sqrt{k}$ yields the last assertion. □

Corollary 1 (A sufficient crowding criterion)**.** *If* $\mathbb{P}(\mathcal{E}_{\delta_n}) \to 1$ *and* $k^{-1}\max_i M_i(\delta_n) \to_p 0$, *then* $\rho_n \to_p 0$. *Under undersmoothing in Theorem 5, paired equivalence follows.*

Proof. Theorem 1(ii) gives $R_n \le \max_i M_i(\delta_n)$ on $\mathcal{E}_{\delta_n}$. The stated convergence follows by division by k. The multiplier in Theorem 6 tends to one under undersmoothing. □

This sufficient criterion is intentionally explicit and can be restrictive. It does not follow from $\delta_n \to 0$ without some control of distance crowding relative to k. In particular, it should not be replaced by an unsupported assertion that close curves have identical neighbors. Theorem 5 explains why lack of paired equivalence does not logically preclude a marginal limit law.

Local Rank Stability and an Explicit Resolution Condition

A global maximum rank displacement is unnecessary: observations carrying zero weight in both estimators have no effect. Define

$$I_n = \{i : r_i \le k \text{ or } \tilde{r}_i \le k\}, \qquad R_n^{\mathrm{loc}} = \max_{i \in I_n} |r_i - \tilde{r}_i|, \qquad \rho_n^{\mathrm{loc}} = R_n^{\mathrm{loc}}/k.$$

The rate below is a sufficient upper bound under two additional local design assumptions: polynomial small-ball behavior and a radial-density bound. It is neither a universal rate for infinite-dimensional predictors nor a necessary reconstruction threshold. In non-polynomial small-ball regimes, the deterministic rank and prediction bounds still apply on $\mathcal{E}_\delta$, but this explicit stochastic rate is not asserted.

Theorem 7 (Local crowding and resolution)**.** *The two bounds in* (18) *hold with* ρ_n^{loc} *in place of* ρ_n*, without the restriction* $R_n \le k$*. The paired order* (19) *and its final implication also hold with this replacement, provided* $\rho_n^{\mathrm{loc}} \to_p 0$*.*

Suppose, in addition, that $F_x(t) \sim c_x t^d$*,* $d \ge 1$*, and that* D_i *has density* $f_x(t) \le Ct^{d-1}$ *for* $0 < t < t_0$*. Let* $\delta_n = o(h_n)$ *and* $\mathbb{P}(\mathcal{E}_{\delta_n}) \to 1$*. Then*

$$\rho_n^{\mathrm{loc}} = O_p\left(\frac{\delta_n}{h_n} + \frac{\log n}{k}\right). \tag{20}$$

Thus $k/\log n \to \infty$*,* $\delta_n = o(h_n)$ *and* $\sqrt{k}h_n^\beta \to 0$ *imply* $\sqrt{k}(\widetilde{m}_k - \widehat{m}_k) \to_p 0$*.*

Proof. The union I_n has at most $2k$ elements, regardless of rank displacement. For every such element, Lipschitz continuity of the profile gives $|\widetilde{w}_i - w_i| \le 2R_n^{\mathrm{loc}}/[k(k+1)]$. Summing this bound and its square over I_n proves the two norm inequalities. The conditional bias and variance argument in Theorem 6 uses only this union and the common localization radius. It therefore proves the claimed replacement, including the convention when $R_n^{\mathrm{loc}} = 0$.

For the rate, work on $D_{(k)} \le h_n$ and $\mathcal{E}_{\delta_n}$; their intersection has probability tending to one. Eventually $\delta_n \le h_n$, and every $i \in I_n$ has $D_i \le 3h_n$. Conditional on $D_i = t \le 3h_n$, the variable $M_i(\delta_n)$ is binomial with $n-1$ trials and success probability

$$p_i = \mathbb{P}\{|D_j - t| \le 2\delta_n\} \le 4C5^{d-1}\delta_n h_n^{d-1} =: v_n,$$

since all relevant distances are in $[0, 5h_n]$. Set $\mu_n = (n-1)v_n$. The binomial Bernstein inequality gives, for every $u > 0$,

$$\mathbb{P}\{M_i > \mu_n + \sqrt{2\mu_n u} + 2u/3 \mid D_i = t\} \le e^{-u}.$$

Integrating over $t \le 3h_n$ preserves this upper bound for the event intersected with $D_i \le 3h_n$. A union bound over all n indices, with $u = 3\log n$, shows that all such counts are at most $\mu_n + \sqrt{6\mu_n \log n} + 2\log n$ except on an event of probability at most n^{-2}. The rank-displacement bound in Theorem 1 now applies to each index of I_n. Since $nh_n^d \asymp k$, division by k gives a bound of order

$$\frac{\delta_n}{h_n} + \sqrt{\frac{\delta_n}{h_n}\frac{\log n}{k}} + \frac{\log n}{k}.$$

The middle term is at most half the sum of the other two terms. This proves (20). Under the final assumptions, $\rho_n^{\mathrm{loc}} = o_p(1)$ and $\sqrt{k}(h_n + \delta_n)^\beta = o(1)$, proving equivalence. □

In the fixed-rank Gaussian reconstruction model below, a sufficient condition for $\delta_n = o(h_n)$ is

$$\frac{p\,(k/n)^{2/d}}{\tau^2 \log(n/\eta_n)} \longrightarrow \infty, \qquad \eta_n \downarrow 0. \tag{21}$$

This is a sufficient resolution condition, not a necessary threshold. For example, $d = 1$, $\beta = 1$, $k \asymp n^{2/5}$, fixed τ, $\eta_n = n^{-1}$ and $p \asymp n^{8/5}$ satisfy all conditions. The density assumption is specific to polynomial small-ball behavior and is not asserted for general infinite-dimensional Gaussian curves.

6. Finite-Sample Inference with Reconstruction Error

For finite-sample coverage, suppose (10) holds globally, with a known exponent β and a known upper bound L, and that response errors are conditionally sub-Gaussian:

$$\mathbb{E}[e^{t\varepsilon_i} \mid \mathcal{G}_n] \le e^{\bar{\sigma}^2 t^2/2}, \qquad t \in \mathbb{R}. \tag{22}$$

Assume a calibrated distance-error bound δ satisfies $\mathbb{P}(\mathcal{E}_\delta) \ge 1-\eta$. A random bound measurable with respect to $\mathcal{G}_n$ is allowed if it has this coverage. The smoothness, conditional response-noise, and reconstruction-error bounds must be valid independently of the response errors used in the interval. Coverage refers to the conditional mean $m(x)$, not a future noisy response. These are certified intervals under stated bounds, not distribution-free intervals.

Theorem 8 (Certified intervals and finite-grid selection)**.** *Let $\mathcal{K} \subset \{2, \dots, n-1\}$ be a fixed finite nonempty set and $0 < \alpha < 1$. Define*

$$A_k = L(\widetilde{D}_{(k)} + \delta)^\beta + \bar{\sigma}\sqrt{2V_k \log(2|\mathcal{K}|/\alpha)}, \qquad k \in \mathcal{K}. \tag{23}$$

Then

$$\mathbb{P}\{m(x) \in [\widehat{m}_k(x) - A_k, \widehat{m}_k(x) + A_k] \text{ for every } k \in \mathcal{K}\} \ge 1 - \alpha - \eta. \tag{24}$$

Consequently, the same coverage bound holds after any measurable selection of k from $\mathcal{K}$. In particular, $\hat{k} = \arg\min_{k\in\mathcal{K}} A_k$, with deterministic tie-breaking, selects the shortest of these certified intervals without invalidating (24).

Proof. Conditional on $\mathcal{G}_n$, independence and (22) imply

$$\mathbb{E}\left[\exp\left(t\sum_i \widetilde{w}_i\varepsilon_i\right)\middle|\mathcal{G}_n\right] \le \exp(t^2\bar{\sigma}^2 V_k/2).$$

Markov's inequality, optimized at $t = u/(\bar{\sigma}^2 V_k)$, bounds the upper tail by $\exp\{-u^2/(2\bar{\sigma}^2 V_k)\}$. Apply the same argument to the negative tail. With $u_k = \bar{\sigma}\sqrt{2V_k \log(2|\mathcal{K}|/\alpha)}$, the two-sided failure probability is at most $\alpha/|\mathcal{K}|$. A union bound over k gives conditional failure probability at most α, regardless of dependence among the estimators.

On $\mathcal{E}_\delta$, Theorem 1(i) gives $D_i \le \widetilde{D}_{(k)} + \delta$ for every positively weighted observation. The absolute conditional bias is therefore at most $L(\widetilde{D}_{(k)} + \delta)^\beta$. Outside the response-error failure event, adding this bias bound to u_k proves all inclusions simultaneously. Unconditionally, their failure is contained in the union of the response-error failure event, of probability at most α, and $\mathcal{E}_\delta^c$, of probability at most η. This proves (24). On the simultaneous event, every selected index is covered, including response-dependent selections. The shortest-interval statement follows directly from minimization over a finite set; it is not an oracle risk inequality. □

For a local Hölder assumption, this result can instead be used only for candidate indices satisfying $\widetilde{D}_{(k)} + \delta \le r_0$. The fixed-grid union bound remains valid for this random subset. If it is empty, the local assumptions certify no interval. Estimating L or $\bar{\sigma}$ from the same sample and treating the estimates as upper bounds without calibration does not preserve the theorem.

6.1. Optimizing the Certified Radius over Nonnegative Weights

The maximum selected distance in (23) bounds all bias terms by the same number. Observation-specific allowances can reduce this conservatism. The scalar-threshold optimization used below follows Anava and Levy [16]; the present statement supplies its reconstruction-adjusted confidence guarantee. We retain a self-contained proof and make no claim that the threshold formula itself is new.

Theorem 9 (Reconstruction-adjusted radius-optimal weights)**.** *Retain the global smoothness, conditional sub-Gaussian, and calibrated reconstruction assumptions. For* $0 < \alpha < 1$ *and* $\bar{\sigma} > 0$*, set*

$$b_i = L(\widetilde{D}_i + \delta)^\beta, \qquad c_\alpha = \bar{\sigma}\sqrt{2\log(2/\alpha)}, \qquad \Delta_n = \{w \in [0,1]^n : \textstyle\sum_i w_i = 1\}.$$

For every $\mathcal{G}_n$*-measurable choice* $w \in \Delta_n$*, the interval centered at* $\sum_i w_i Y_i$ *with half-width*

$$A(w) = \sum_i b_i w_i + c_\alpha \|w\|_2 \tag{25}$$

has coverage at least $1 - \alpha - \eta$ *for* $m(x)$*. Let* $t > \min_i b_i$ *be the unique solution of*

$$\sum_i (t - b_i)_+^2 = c_\alpha^2. \tag{26}$$

Then the unique minimizer of A *on* Δ_n *is*

$$w_i^* = \frac{(t - b_i)_+}{\sum_j (t - b_j)_+}, \qquad \min_{w \in \Delta_n} A(w) = t. \tag{27}$$

In particular, t *is no greater than the singleton-grid radius in* (23) *for any* k*. Selection of* w^* *requires no multiplicity penalty because it uses predictors and calibrated constants only.*

Proof. On $\mathcal{E}_\delta$, $D_i \le \widetilde{D}_i + \delta$ for all i, so the conditional bias has absolute value at most $\sum_i b_i w_i$. Conditional independence bounds the moment-generating function of $\sum_i w_i \varepsilon_i$ by $\exp(\bar{\sigma}^2 \|w\|_2^2 u^2/2)$. The two-sided Chernoff argument gives a conditional probability at most α that its absolute value exceeds $c_\alpha \|w\|_2$. Integrating and adding the reconstruction failure probability η proves coverage, including for a random predictor-measurable vector w.

The left side of (26) is continuous, zero at $\min_i b_i$, strictly increasing above this point, and tends to infinity. Its root therefore exists and is unique. Put $v_i = (t - b_i)_+$ and $S = \sum_i v_i > 0$. The equation gives $\|v\|_2 = c_\alpha$. For every $w \in \Delta_n$, since $t - b_i \le v_i$,

$$t - \sum_i b_i w_i = \sum_i (t - b_i) w_i \le \sum_i v_i w_i \le c_\alpha \|w\|_2.$$

Thus $A(w) \ge t$. At $w = v/S$, equality holds in both inequalities and $A(v/S) = t$. Equality in Cauchy–Schwarz requires w proportional to v; normalization forces the unique choice in (27). This proves both optimality and uniqueness without an unstated regularity condition on the active set.

For triangular rank weights, $\sum_i b_i \widetilde{w}_i \le L(\widetilde{D}_{(k)} + \delta)^\beta$ and $\|\widetilde{w}\|_2 = \sqrt{V_k}$. Comparing this feasible vector with the minimum proves the radius comparison. Conditional on $\mathcal{G}_n$, the selected weights are fixed, so optimizing over a continuum of weights does not require simultaneous control of response-dependent choices. □

Corollary 2 (Reconstruction floor and rates)**.** *The optimized half-width satisfies*

$$L\delta^\beta + \frac{c_\alpha}{\sqrt{n}} \le t \le L(\widetilde{D}_{(k)} + \delta)^\beta + c_\alpha \sqrt{V_k} \quad (2 \le k < n).$$

For the following asymptotic statements, let $\delta_n \to 0$ *be deterministic, assume* $\mathbb{P}(\mathcal{E}_{\delta_n}) \to 1$*, and retain the small-ball localization assumptions. For fixed* α*,* $t = O_p((h_n + \delta_n)^\beta + k^{-1/2})$*. In the polynomial regime, the balanced choice of* k *yields* $t = O_p(n^{-\beta/(2\beta+d)} + \delta_n^\beta)$*.*

Proof. Each $b_i \ge L\delta^\beta$ and Cauchy–Schwarz gives $\|w\|_2 \ge n^{-1/2}$ for every simplex vector. Minimize to obtain the lower bound. The upper bound is Theorem 9. On the high-probability localization event, $\widetilde{D}_{(k)} \le h_n + \delta_n$; insert this into the upper bound. Since $0 < \beta \le 1$, the inequality $(a + b)^\beta \le a^\beta + b^\beta$ and the balanced k from Theorem 3 give the final rate. □

For fixed reconstructed distances and $\beta = 1$, the common tolerance contributes only an additive term: $A_\delta(w) = A_0(w) + L\delta$. Thus the optimal weights do not depend on δ alone, and the optimal half-width increases by $L\delta$.

The lower bound is a floor for this certified objective, not an information-theoretic lower bound for arbitrary procedures. The rate comparison uses any admissible k as a witness: the optimizer itself does not require d or an explicit choice of k. It still requires valid L, β, $\bar{\sigma}$ and δ. In particular, optimizing weights cannot eliminate conservatism due to a large reconstruction bound.

Numerical threshold and reported radius

All interval generators use Brent's method on $[\min_i b_i, \min_i b_i + c_\alpha]$, with absolute root tolerance 10^{-12} and relative tolerance $4\epsilon_{\text{mach}}$ (approximately 8.88×10^{-16} in float64). After normalizing the nonnegative threshold weights, the reported half-width is evaluated directly as $A(w) = b^\top w + c_\alpha \|w\|_2$, rather than substituting the approximate root. Thus, the stated bias and tail inequality applies to the actual feasible numerical weights; numerical optimality is approximate.

6.2. A Calibrated Reconstruction Example

Consider the finite-rank model $X_i(t) = \sum_{j=1}^{J} a_{ij}\phi_j(t)$ with $\phi_j(t) = \sqrt{2}\sin(\pi j t)$ and independent observation errors $N(0, \tau^2)$ at $t_\ell = \ell/(p+1)$, $\ell = 1, \ldots, p$, where $J \le p$. Orthogonality on this grid gives $\Phi^\top \Phi = (p+1)I_J$. Least squares yields

$$\tilde{a}_i - a_i \sim N_J(0, \tau^2 I_J/(p+1)). \tag{28}$$

The same experiment is used for the target independently. Let $Q_J(u)$ be the u-quantile of χ^2_J. A union bound over the $n+1$ reconstructions gives

$$\delta = \frac{2\tau}{\sqrt{p+1}}\sqrt{Q_J\Big(1 - \frac{\eta}{n+1}\Big)}, \qquad \mathbb{P}(\mathcal{E}_\delta) \ge 1 - \eta. \tag{29}$$

Indeed, each coefficient-error norm exceeds half this δ with probability $\eta/(n+1)$; the orthonormal basis equates coefficient and functional norms, and the triangle inequality gives (4).

For fixed J and fixed η, the bound is of order $\tau\sqrt{\log n/p}$. For asymptotic high-probability events, one may take $\eta_n \downarrow 0$ slowly, adding $\log(1/\eta_n)$ to the order. A sufficient resolution requirement for bias-free inference is then

$$\sqrt{k}\left[h_n + \tau\sqrt{\frac{\log(n/\eta_n)}{p}}\right]^\beta \longrightarrow 0.$$

With $\beta = 1$, $p/(k\log(n/\eta_n)) \to \infty$ makes the reconstruction part negligible for fixed τ. General infinite-rank curves require an additional approximation-bias bound; (29) must not be applied to them without it.

6.3. Calibration Status and Sensitivity

Theorems 8 and 9 are finite-sample guarantees only when L, $\bar{\sigma}$, and δ are valid upper bounds independently of the response noise used in the interval. This condition is substantive. In the simulation model, L and $\bar{\sigma}$ are known from the data-generating mechanism, so those experiments are theoretical benchmarks rather than a fully data-adaptive procedure.

The reconstruction constant can nevertheless be calibrated from independent measurement-error information. In model (28), if τ is unknown but an independent calibration experiment yields an upper confidence bound τ_U satisfying $\Pr(\tau \le \tau_U) \ge 1 - \xi$,

then replacing τ by τ_U in (29) gives $\Pr(\mathcal{E}_\delta) \geq 1-\eta-\xi$ by a union bound. The corresponding interval coverage is at least $1-\alpha-\eta-\xi$. Here the conditioning information is enlarged to include the independent calibration experiment. For example, with M independent zero-mean Gaussian calibration errors and sample second moment $\widehat{\tau}^2 = M^{-1}\sum_{j=1}^{M} e_j^2$, one may take $\tau_U^2 = M\widehat{\tau}^2/\chi^2_{M,\xi}$, where $\chi^2_{M,\xi}$ is the lower ξ quantile. An analogous scale upper bound is available for homoscedastic Gaussian response noise when independent residual calibration data exist. By contrast, under the paper's general heteroscedastic model there is no distribution-free way to turn an ordinary fitted residual scale into a uniform conditional sub-Gaussian bound $\bar{\sigma}$. Likewise, a globally valid Hölder constant L cannot be certified from the same finite sample without additional structural assumptions. In applications lacking external or independent bounds, the finite-sample intervals should therefore be interpreted as benchmark certificates or sensitivity analyses, not deployment-ready confidence procedures.

7. Monte Carlo Study

7.1. Design, Targets, and Distinct Inferential Levels

The baseline factorial simulation has three scenarios, $n \in \{200, 500, 1000\}$, $p \in \{16, 32, 64\}$, and 400 independent replications per cell, giving 10,800 datasets. Streams use `SeedSequence([20260917, scenario index, n, p])`. Each curve has six sine coefficients. The fixed target is the exact zero curve, with its own independent noisy reconstruction, and $m(0) = 0$ in all scenarios.

- S1: $a_{ij} = j^{-2}Z_{ij}$ with independent $Z_{ij} \sim N(0,1)$; $m(a) = \sin(a_1) + 0.5\sin(a_2)$; response-noise standard deviation 0.35; measurement-noise standard deviation $\tau = 0.15$.
- S2: $a_{ij} = j^{-1}Z_{ij}$; $m(a) = \sin(a_1) + 0.3\{1 - \cos(a_2)\}$; response-noise standard deviation $0.3 + 0.2/(1 + e^{-a_1})$; $\tau = 0.6$.
- S3: coefficients as in S2, with an independent equiprobable -1 or $+1$ shift in a_{i1}; $m(a) = \sin(a_1) + 0.5\sin(a_2)$; response-noise standard deviation 0.35; $\tau = 0.6$.

Response errors are conditionally Gaussian and independent of measurement errors. Baseline coefficient errors are generated from the exact least-squares law (28); no grid simulation is needed for that law. These finite-rank experiments do not demonstrate a general infinite-dimensional rate.

Predictive comparisons below are separated from diagnostics of reconstruction perturbation, mean-interval coverage, and asymptotic resolution. Ordinary normal intervals use the nominal level 0.95, namely multiplier $z_{0.975} \simeq 1.96$, independently of the certified-interval allocation. Certified intervals use $\alpha = 0.04$ for the response tail and $\eta = 0.01$ for reconstruction failure, giving the lower coverage bound 0.95 only under their valid bounds. We set $\beta = 1$, $L = \sqrt{1.25}$ in S1/S3 and $L = \sqrt{1.09}$ in S2, using gradient-norm bounds, and $\bar{\sigma} = 0.35$ in S1/S3 and 0.5 in S2. These are known generating constants, not fitted response-based estimates.

Monte Carlo uncertainty uses independent training replications. We report standard errors of MSE, interval width, and coverage in the archived cellwise CSV files. RMSE standard errors in target comparisons use the delta method, $\mathrm{SE}(\widehat{\mathrm{MSE}})/(2\widehat{\mathrm{RMSE}})$. For pooled scenario comparisons, equally weighted scenario means are combined, and their independent variance contributions are added. Evaluating several methods or targets on the same training dataset does not increase the independent sample size.

7.2. Predictive Weighting Comparisons

The baseline fixes $k = \lfloor\sqrt{n}\rfloor$ for triangular rank weights, uniform kNN, and triangular distance weights $(1 - \widetilde{D}_i/\widetilde{D}_{(k+1)})_+$ on the first k neighbors. All use the same reconstructed

distances. This isolates weighting effects and does not compare independently tuned methods. Figure 1 reports RMSE against n; Table 6 gives the $n = 500$, $p = 32$ comparison. Complete factorial summaries include MSE uncertainty, MAE, bias, and replication records in S1.

Table 6. Predictive performance at $n = 500$, $p = 32$, with the common fixed $k = 22$. Each row has 400 independent training replications; SE(MSE) is a Monte Carlo standard error.

Scenario	Method	RMSE	SE(MSE)	MAE	Bias
S1	Distance taper	0.104	0.0009	0.082	−0.004
S1	Rank kNN	0.096	0.0007	0.077	−0.002
S1	Uniform kNN	0.083	0.0005	0.068	0.002
S2	Distance taper	0.150	0.0016	0.120	0.019
S2	Rank kNN	0.144	0.0015	0.115	0.020
S2	Uniform kNN	0.138	0.0013	0.111	0.020
S3	Distance taper	0.154	0.0016	0.125	0.006
S3	Rank kNN	0.149	0.0015	0.121	0.004
S3	Uniform kNN	0.142	0.0014	0.114	0.003

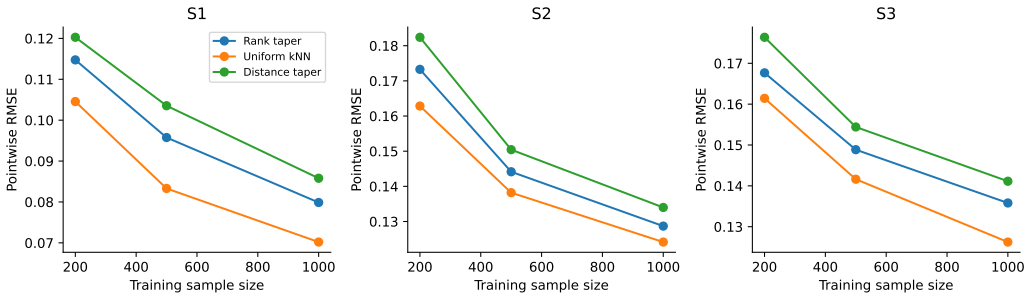


Figure 1. Pointwise RMSE of three weighting schemes at $p = 32$, using the same k. Rankings in these scenarios do not establish uniform predictive dominance.

7.3. Target Location, Training-Only Tuning, and Paired Comparisons

An independent target experiment uses $n = 500$, $p = 32$, and 45 fresh training samples per scenario. Each training sample has three targets: zero (center), one independent design draw, and the independent candidate nearest the 95th percentile of coefficient norm among 300 draws (tail). Each target has independent reconstruction noise. There are 135 independent training samples and 405 target evaluations per method.

All three neighbor methods select $k \in \{5, 10, 15, 20, 30, 40\}$ by four-fold cross-validation on the reconstructed training sample. The Gaussian functional Nadaraya–Watson comparator uses bandwidths at the 4th, 8th, 15th, 25th, and 40th percentiles of pairwise outer-training L^2 coefficient distances. PCR uses component counts $\{2, 3, 4, 6\}$ and ridge penalties $\{0.01, 0.1, 1, 10\}$. PCA includes mean centering and is refitted separately on each validation training fold, then on the complete training sample; responses are centered internally by ridge regression in the corresponding training fit. No component standardization is applied. The kernel bandwidth grid uses outer-training predictors only and is fixed across its validation folds; it never uses target data. Family-specific CV seeds are fixed and reported in the code.

The kernel percentile rule and neighbor grid are the same in the simulation and Tecator, with RMS distances in Tecator. PCR grids differ: Tecator uses $\{3, 6, 10, 15\} \cap \{1, \ldots, p\}$ components and $\{0.001, 0.01, 0.1, 1, 10\}$ penalties, whereas the simulation has only six reconstructed coefficient directions. Thus, predictive rankings cannot be attributed solely to the regression rule independently of design, target location, representation, and tuning. No target response or test response enters tuning.

Table 7 presents target-location comparisons.

Table 7. Target comparisons pooled with equal scenario weights. There are 45 independent training replications per scenario. SE columns are Monte Carlo standard errors; RMSE uncertainty uses the delta method.

Target	Method	RMSE	SE(RMSE)	MAE	SE(MAE)
Center	Rank kNN	0.111	0.006	0.088	0.006
Center	Uniform kNN	0.114	0.007	0.092	0.006
Center	Distance taper	0.112	0.006	0.089	0.006
Center	Functional kernel	0.067	0.004	0.053	0.003
Center	PCR	0.059	0.004	0.046	0.003
Random	Rank kNN	0.130	0.009	0.099	0.007
Random	Uniform kNN	0.132	0.008	0.104	0.007
Random	Distance taper	0.129	0.008	0.100	0.007
Random	Functional kernel	0.187	0.008	0.155	0.008
Random	PCR	0.259	0.015	0.215	0.012
Tail	Rank kNN	0.164	0.012	0.120	0.008
Tail	Uniform kNN	0.173	0.013	0.125	0.009
Tail	Distance taper	0.165	0.012	0.121	0.009
Tail	Functional kernel	0.222	0.014	0.161	0.010
Tail	PCR	0.437	0.018	0.372	0.014

At the center, PCR and functional kernel regression have the smallest RMSEs (0.059 and 0.067), versus 0.111 for rank kNN. At random targets, rank kNN has RMSE 0.130, compared with 0.187 for the kernel and 0.259 for PCR; at tail targets the corresponding values are 0.164, 0.222 and 0.437. These are design-specific results, not a universal superiority claim. Scenario-specific results with uncertainty are in Supplementary Analyses, Table S1.

Table 8 presents paired predictive comparisons.

Table 8. Paired differences in squared prediction error, first method minus second, pooled over scenarios at each target. Negative values favor the first method. Pairing is retained within each independent training replication; MCSE is not based on treating methods or targets as independent observations.

Target	Comparison	Mean MSE Difference	MCSE
Center	Rank-Uniform	−0.00069	0.00072
Random	Rank-Uniform	−0.00067	0.00092
Tail	Rank-Uniform	−0.00312	0.00148
Center	Rank-Distance	−0.00025	0.00035
Random	Rank-Distance	0.00028	0.00048
Tail	Rank-Distance	−0.00038	0.00071
Center	Uniform-Distance	0.00044	0.00086
Random	Uniform-Distance	0.00095	0.00100
Tail	Uniform-Distance	0.00273	0.00160

The small gaps between rank and distance tapering are uncertain: the pooled mean paired MSE differences are −0.00025, 0.00028, and −0.00038 at center, random, and tail targets, with MCSEs 0.00035, 0.00048, and 0.00071. They provide no stable evidence of an advantage of rank tapering over distance tapering. For the across-target comparison, squared errors are first averaged across the three targets within each training replication; complete scenario-specific paired differences appear in Supplementary Analyses, Table S2. The method's additional value is its explicit rank normalization and audit of reconstruction-induced changes, which do not ensure better prediction than a suitable global representation.

7.4. Reconstruction Perturbation and Inferential Diagnostics

The paired difference $|\widetilde{m}_k - \widehat{m}_k|$ uses the same responses and the same k and therefore isolates reconstruction effects. It is distinct from prediction error relative to $m(x)$. Figure 2 reports this difference and ordinary mean-interval coverage; neither is evidence for a universal rank-stability rate.

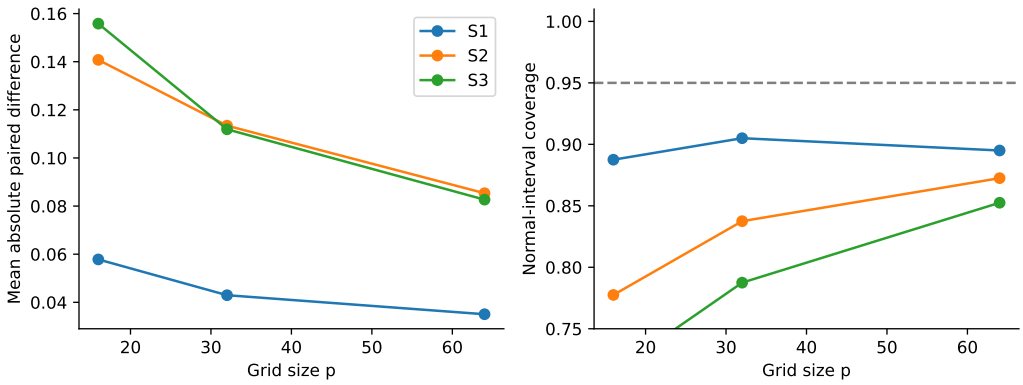


Figure 2. Theoretical diagnostics at $n = 500$. (**Left**): mean absolute prediction perturbation on the same sample. (**Right**): ordinary nominal 95% studentized mean-interval coverage, without finite-sample certification. The gray dashed horizontal line in the right panel marks the nominal 95% coverage level. Fixed noisy resolution need not satisfy vanishing reconstruction error or undersmoothing.

Table 9 presents mean-interval diagnostics.

Table 9. Mean-interval diagnostics at $p = 32$, common fixed k, and 400 replications per row. N is the ordinary nominal 95% interval. C is certified to have coverage of at least 95% only under independently valid L, $\bar{\sigma}$, and δ, with $(\alpha, \eta) = (0.04, 0.01)$. Widths are total widths.

Scenario	n	Cov. N	SE(Cov. N)	Width N	Cov. C	Width C	SE(Width C)
S1	200	0.897	0.015	0.411	1.000	1.78	0.003
S1	500	0.905	0.015	0.330	1.000	1.60	0.002
S1	1000	0.907	0.015	0.280	1.000	1.50	0.001
S2	200	0.833	0.019	0.502	1.000	4.37	0.004
S2	500	0.838	0.018	0.398	1.000	4.17	0.003
S2	1000	0.825	0.019	0.346	1.000	4.05	0.002
S3	200	0.848	0.018	0.491	1.000	4.52	0.005
S3	500	0.787	0.020	0.384	1.000	4.35	0.004
S3	1000	0.767	0.021	0.328	1.000	4.23	0.003

Ordinary coverage remains below 0.95; in S3 it decreases from 0.8475 to 0.7675 as n rises and widths shrink. Fixed noisy resolution does not satisfy the vanishing-error condition for bias-free inference at the latent zero curve. The certified intervals cover in all 400 replications of each displayed cell but are much wider, especially in S2/S3. For 400 successes, the two-sided exact 95% binomial interval for the experiment-specific coverage probability is approximately $[0.991, 1]$; this describes simulation uncertainty and does not verify the assumed bounds in an application. At coverage 0.95, the nominal binomial MCSE is about 0.011. Coverage equal to one is evidence of conservatism in these experiments, not evidence of practical calibration or efficient inference.

7.5. Radius Optimization and Calibration Sensitivity

The separate interval experiment has the same 27 cells and 200 fresh replications per cell (5400 datasets), using `SeedSequence([920260919, scenario index, n, p])`. It compares the original maximum-radius certificate with $A(w)$ at the original rank weights

and radius-optimal weights. Generating constants and $(\alpha, \eta) = (0.04, 0.01)$ are unchanged. Reported optimized half-widths are directly evaluated as $A(w)$ for the numerical weights.

Table 10 presents certified interval-width comparisons.

Table 10. Certified mean intervals at $n = 500$, $p = 32$, 200 replications per scenario. Ratio is the mean paired width ratio to the maximum-radius certificate. Coverage guarantees require independently valid smoothness, noise, and reconstruction bounds.

Scenario	Certificate	Coverage	Width	Ratio	RMSE
S1	max radius	1.000	1.604	1.000	0.095
S1	optimized	1.000	1.384	0.863	0.060
S1	weighted bias	1.000	1.467	0.915	0.095
S2	max radius	1.000	4.175	1.000	0.138
S2	optimized	1.000	3.839	0.919	0.116
S2	weighted bias	1.000	3.932	0.942	0.138
S3	max radius	1.000	4.335	1.000	0.136
S3	optimized	1.000	4.022	0.928	0.129
S3	weighted bias	1.000	4.038	0.932	0.136

Figure 3 illustrates the paired width ratios. Optimization reduces mean paired widths by 13.7%, 8.1%, and 7.2% in S1–S3; the radius ordering holds in every generated dataset. Nevertheless, all displayed coverages equal one, and the intervals remain wide. Optimization of this bound does not establish practical efficiency or a general prediction advantage. Its center uses known smoothness and noise bounds, unlike the tuned predictive baselines.

We quantify sensitivity at $n = 500$, $p = 32$ using 120 fresh replications in each of S1–S3. One constant at a time is multiplied by 0.75, 1, 1.25, or 1.5, while the other two remain at their generating values. Factors below one are diagnostic stress tests and do not retain the theorem's guarantee; factors at or above one do if the baseline constant is valid. Table 11 shows that all tested intervals remain empirically conservative, while width changes materially with the supplied constants. Thus the numerical evidence makes the dependence on calibration explicit. In the δ sensitivity rows, $\beta = 1$ and the reconstructed distances are fixed within each replication; only the interval width changes, in agreement with $A_\delta(w) = A_0(w) + L\delta$.

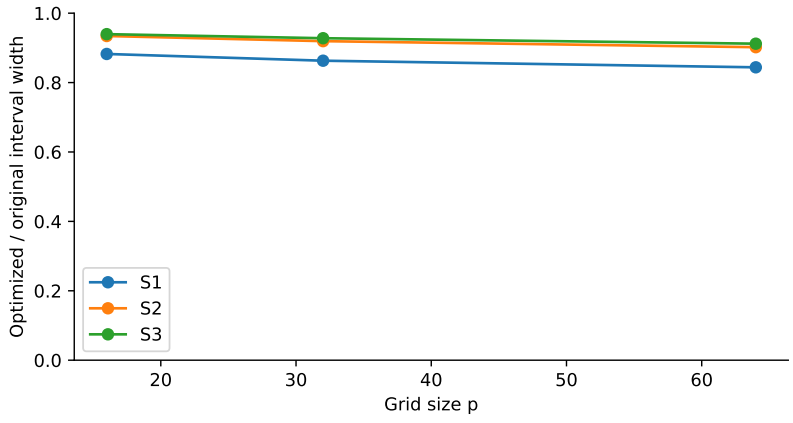


Figure 3. Mean paired optimized-to-original certified width ratio at $n = 500$, 200 replications per point. A ratio no greater than one concerns optimization of the specified bound. Finite-sample coverage still requires independently valid L, $\bar{\sigma}$, and δ.

Table 11. Sensitivity of the optimized certificate to one-at-a-time multiplicative changes in L, $\bar{\sigma}$, and δ at $n = 500$, $p = 32$, aggregated over 120 replications in each of S1–S3. Coverage is for $m(0)$ and width is the total interval width. Finite-sample coverage is guaranteed only for independently valid upper bounds; multipliers below one are unguaranteed stress tests.

Constant Varied	Multiplier	Coverage	Mean Width	RMSE
L	0.75	1.000	2.398	0.109
L	1.00	1.000	3.082	0.115
L	1.25	1.000	3.753	0.121
L	1.50	1.000	4.414	0.126
δ	0.75	1.000	2.605	0.115
δ	1.00	1.000	3.082	0.115
δ	1.25	1.000	3.560	0.115
δ	1.50	1.000	4.037	0.115
$\bar{\sigma}$	0.75	1.000	2.981	0.123
$\bar{\sigma}$	1.00	1.000	3.082	0.115
$\bar{\sigma}$	1.25	1.000	3.170	0.110
$\bar{\sigma}$	1.50	1.000	3.249	0.107

7.6. Sensitivity to Sample Size and Neighborhood Size

To separate response-based tuning from fixed-neighborhood behavior, we replay the 3600 baseline datasets at $p = 32$ in S1–S3 and $n \in \{200, 500, 1000\}$. On each dataset, we evaluate $k \in \{5, 10, 20, 40, \lfloor \sqrt{n} \rfloor\}$. These are additional evaluations of existing independent datasets, not new datasets. MSE uncertainty, ordinary and certified coverage, and both widths are recorded for all 45 cells in Supplementary Analyses, Table S3, and replication CSV files. Each fixed-k certificate uses a singleton grid; these marginal comparisons do not claim simultaneous coverage after selecting among the five displayed intervals.

Table 12 presents neighborhood-size sensitivity results.

At $n = 1000$ in S3, increasing k from 5 to 40 reduces RMSE from 0.220 to 0.129 and ordinary width from 0.610 to 0.295, but ordinary coverage changes from 0.800 to 0.755. Certified widths remain approximately 4.44 and 4.26, with empirical coverage at one. Better point prediction or a narrower normal interval therefore does not establish valid mean inference. The direction of width changes also depends on the balance between bias and variance; fixed noisy reconstruction supplies a nonvanishing contribution.

Table 12. Neighborhood sensitivity in S3 at $p = 32$, 400 paired replications per row; the full three-scenario results are in S1. N denotes ordinary nominal 95% mean intervals; C denotes intervals certified only under the independently valid generating bounds.

n	k	RMSE	Cov. N	Width N	Cov. C	Width C
200	5	0.230	0.787	0.648	1.000	4.59
200	40	0.136	0.800	0.341	1.000	4.81
500	5	0.223	0.782	0.630	1.000	4.50
500	40	0.132	0.745	0.305	1.000	4.44
1000	5	0.220	0.800	0.610	1.000	4.44
1000	40	0.129	0.755	0.295	1.000	4.26

7.7. Alternative Reconstruction Mechanisms

We add three independently seeded reconstruction mechanisms at $n = 500$, $p = 32$ with 120 replications in each scenario–mechanism cell, giving 1080 new datasets. Seeds are `SeedSequence([20261012,3,scenario index,mechanism index,rep])`. Responses follow the same means and Gaussian noise models; measurement information remains independent of response errors.

1. **Irregular grids, Gaussian measurement errors.** Each dataset has a shared sorted grid of 32 independent uniform points in $[0.02, 0.98]$. Six-component least squares has covariance $\tau^2(\Phi^\top\Phi)^{-1}$. We use its largest eigenvalue to obtain a conditional Gaussian norm bound; ill-conditioned grids increase the bound rather than being discarded.
2. **Regular grids, Laplace measurement errors.** We explicitly generate independent Laplace errors with scale $\tau/\sqrt{2}$ at the regular measurement grid and project them by least squares. Their variance is τ^2. A coordinatewise sub-gamma inequality and a union bound give a valid reconstruction allowance; the Gaussian chi-square formula is not used for these errors.
3. **Truncated population principal components.** The latent predictor has 12 sine coefficients with the scenario-specific decay, but reconstruction retains only the first six population principal-component directions. We combine a Gaussian retained-coefficient error bound with a bound on the omitted Gaussian tail. This is a known population-eigenbasis benchmark, not an estimated-FPCA procedure; it does not include uncertainty from learning eigenfunctions.

The preceding construction specifies the reconstruction bounds used in these stress tests. All upper bounds are supplied by the measurement/design mechanism, independently of the responses. The finite-sample certification arguments use the error event and therefore apply beyond Gaussian measurement errors when a valid event bound is available; the special Gaussian resolution formula does not automatically transfer. These additional experiments remain finite-rank stress tests and do not validate an arbitrary infinite-dimensional reconstruction model.

Table 13 presents alternative reconstruction stress-test results.

All displayed certified coverages equal one; with 120 successes, the exact two-sided 95% binomial lower limit is approximately 0.970. Ordinary coverage ranges from 0.750 to 0.883. Conservatism becomes more severe for the coordinatewise Laplace bound: mean optimized widths reach 9.36 and 9.95 in S2/S3, compared with retrospective-reference widths 3.05 and 3.18. Thus, validity under a broader measurement law does not imply a useful interval width. Truncation and irregular designs also require additional bound information. Separate deterministic perturbation certificates are evaluated on all 1080 datasets, and joint certificates on the first 12 replications per cell (108 datasets); all contain the realized perturbations on their error events. Counts, realized slack, and uncertainty are supplied in S1, not interpreted as proof of prospective usefulness.

Table 13. Alternative reconstructions: 120 independent datasets per row. RMSE and Cov. N refer to fixed rank $k = 22$; Cov. C and Width C refer to radius-optimal certified mean intervals under independently valid generating bounds. Ref. width uses a retrospective error bound computed from latent curves and is not prospectively available.

Scenario	Reconstruction	RMSE	Cov. N	Cov. C	Width C	Ref. Width
S1	Irregular Gaussian	0.101	0.883	1.000	1.72	1.20
S1	Regular Laplace	0.089	0.883	1.000	2.86	1.16
S1	Truncated PCs	0.097	0.867	1.000	1.62	1.15
S2	Irregular Gaussian	0.146	0.783	1.000	5.02	3.17
S2	Regular Laplace	0.161	0.767	1.000	9.36	3.05
S2	Truncated PCs	0.151	0.817	1.000	6.11	3.20
S3	Irregular Gaussian	0.136	0.833	1.000	5.48	3.39
S3	Regular Laplace	0.142	0.850	1.000	9.95	3.18
S3	Truncated PCs	0.165	0.750	1.000	6.46	3.38

7.8. Resolution Diagnostics with an Exact-Curve Comparator

A one-coefficient experiment isolates the sufficient resolution condition: $X(t) = a\phi_1(t)$ with $a \sim N(0,1)$, $m(a) = \sin(a)$, response errors $N(0, 0.35^2)$, measurement noise $\tau = 0.6$, and a zero target with independent reconstruction noise. We use $k = \lfloor n^{2/5} \rfloor$, $n \in \{200, 800, 2000\}$, and either $p = 32$ or $p = \lceil n^{8/5} \rceil$, with 400 replications per cell (2400 datasets). Streams use `SeedSequence([197160,n,p])`. The exact-curve and reconstructed studentized intervals use the same responses, target, and k on every replication.

Table 14 presents resolution diagnostics.

Table 14. Resolution diagnostics with nominal 95% ordinary mean intervals. Cov. R uses reconstructed curves and Cov. E exact curves on the same replications. Gap is mean $\sqrt{k}|\widetilde{m}_k - \widehat{m}_k|$ and displacement is mean ρ_n^{loc}.

Resolution	n	p	k	Cov. R	Cov. E	Gap	Displacement
fixed	200	32	8	0.802	0.843	0.484	5.017
increasing	200	4805	8	0.875	0.882	0.129	0.511
fixed	800	32	14	0.815	0.938	0.516	12.815
increasing	800	44152	14	0.925	0.917	0.090	0.410
fixed	2000	32	20	0.730	0.925	0.613	24.569
increasing	2000	191271	20	0.930	0.922	0.086	0.369

With increasing resolution, mean scaled gaps fall from 0.129 to 0.086, and reconstructed coverage rises from 0.875 to 0.930. Exact-curve coverage is 0.883, 0.918, and 0.923: appreciable finite-sample undercoverage remains even without reconstruction, consistent with small neighborhoods and studentization error. These close coverage differences should not be overinterpreted. Under fixed resolution at $n = 2000$, reconstructed coverage is 0.730 compared with 0.925 for exact curves, exposing an additional reconstruction effect. This low-dimensional diagnostic supports the direction of the sufficient asymptotic conditions, not their necessity or general nominal coverage at finite sample sizes.

8. Real-Data Application: Tecator Spectra

8.1. Data Provenance and Leakage Prevention

The Tecator data were obtained from the public CRAN mirror of the `fda.usc` package [19–21]. The archived R object contains 215 spectra, each observed at 100 wavelengths spanning 850–1050 nm, and reference fat, water, and protein contents determined by analytic chemistry. We use fat percentage as the response. The source R object, its documentation, a CSV conversion, acquisition URLs, and cryptographic checksums are supplied with the analysis.

Figure 4 illustrates representative measured spectra and the held-out rank-smoother predictions. There are 193 distinct absorbance vectors: 171 occur once and 22 occur twice. We group identical spectra before splitting and average their fat labels. This prevents exact spectrum duplicates from appearing in both training and test samples. It defines a benchmark over distinct spectra; it does not establish independence between all biological specimens, for which additional identifiers are unavailable. The mapping from original rows to the grouped data is included.

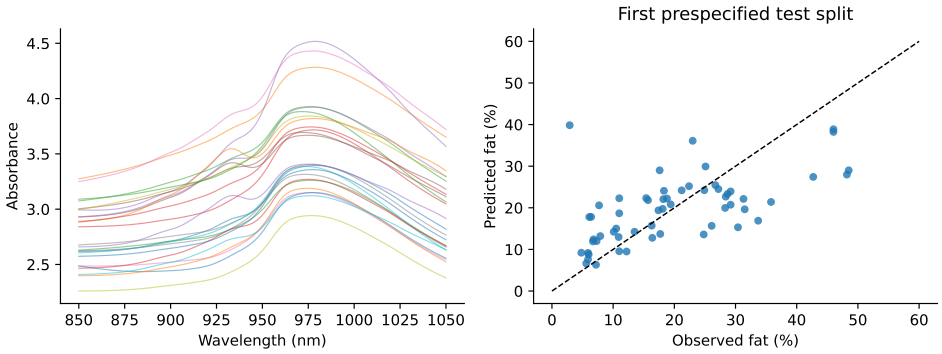


Figure 4. Tecator application. (**Left**): 25 deterministically selected distinct spectra; differently colored curves distinguish individual spectra and do not denote separate experimental groups. (**Right**): observed and predicted fat content for rank smoothing on the first prespecified held-out split at full resolution. The black dashed diagonal indicates perfect prediction; points represent observed–predicted pairs.

8.2. Training, Tuning, and Resolution Perturbations

Fifty reproducible random splits allocate 135 distinct spectra to training and 58 to testing. Split r uses seed $20260917 + r$, $r = 0, \ldots, 49$. For each split, we retain $p \in \{10, 20, 50, 100\}$ approximately equally spaced wavelength indices, including endpoints, and interpolate linearly back to the original 100-point grid. No noise is added to the real spectra. Here, the complete measured 100-point spectrum is the reference, not an unobserved continuous noiseless curve.

The distance is the root-mean-square difference between absorbance vectors. All three neighbor methods select $k \in \{5, 10, 15, 20, 30, 40\}$ by four-fold cross-validation entirely within the training sample. Ties in validation loss select the first candidate. A functional Gaussian-kernel Nadaraya–Watson comparator selects its bandwidth from the 4th, 8th, 15th, 25th and 40th percentiles of pairwise training distances, using the same four folds. The grid uses training predictors only; no test spectrum enters bandwidth selection. A principal-component ridge regression (PCR) baseline selects components from $\{3, 6, 10, 15\} \cap \{1, \ldots, p\}$ and the ridge penalty from $\{0.001, 0.01, 0.1, 1, 10\}$. Excluding components above p avoids redundant null directions after interpolation from p retained wavelengths. Candidate PCR losses within 10^{-12} of the minimum are treated as tied, selecting the first candidate in the documented grid order. PCA is fitted separately inside each validation training fold and refitted on the complete outer training set. No test outcomes enter tuning, centering, PCA fitting, or reconstruction. Principal components provide predictive compression; their individual axes are not assigned a physical interpretation. This use is consistent with the interpretability concerns discussed by Vilenchik et al. [18].

For perturbation auditing, the reduced-resolution rank prediction is compared with the full-resolution rank prediction at the *same* selected k. Otherwise, parameter-selection changes would be confounded with rank changes. We calculate a common split-wide allowance $\delta_{\text{split}} = e_{\text{tr,max}} + e_{\text{te,max}}$, where $e_i = \|\widetilde{X}_i - X_i\|_{\text{RMS}}$, as a single conservative tolerance for every held-out target. It uses only predictor reconstruction differences, not test responses. A fixed target x_j also admits the smaller valid reference allowance $\delta_j = e_{\text{tr,max}} + e_j$. The comparison below quantifies the extra conservatism caused by the common test-split maximum. The use of full spectra for this diagnostic is deliberate: it is a controlled coarsening experiment. It is not a deployment certificate when the reference spectrum is unavailable.

8.3. Predictive Results and Scope of the Stability Audit

This study is a stability audit under a specified linear-interpolation coarsening mechanism applied to measured spectra. It does not validate inference for an unobserved continuous noiseless spectrum or identify its reconstruction error. Predictive performance and perturbation certification answer separate questions. At full resolution, the average test RMSE is 8.584 for rank smoothing, 9.055 for uniform kNN, 8.438 for distance tapering, 10.507 for functional kernel regression, and 2.983 for PCR (Table 15). PCR has the smallest mean RMSE among the five tested methods. The real application does not support a claim of state-of-the-art fat-content prediction by rank weighting. Monotone distance-scale invariance does not correct an insufficiently informative neighborhood geometry.

For rank smoothing, the mean absolute difference from the full-resolution prediction decreases from 0.604 percentage points at $p = 10$ to 0.175 at $p = 20$ and 0.007 at $p = 50$. Mean neighbor-set overlap rises from 0.969 to 0.992 and 0.999, respectively. These calculations use the same k within each paired comparison. All computed deterministic prediction certificates contain their corresponding paired differences. However, the certificates are often loose, reflecting both the maximum reconstruction-error bound and the independent rank-interval relaxation. Their role is an auditable upper bound rather than a sharp empirical error estimate. A realized difference near zero can occur because ranks do not change or because response-weight changes cancel. The certificate maximizes over every permitted rank permutation, including unrealized reorderings that need not exhibit that cancellation. A worst-case upper bound can therefore stay large when the realized difference is zero; split-wide reconstruction maxima enlarge that uncertainty further.

Table 15. Tecator held-out performance over 50 splits. RMSE and MAE are in fat percentage points. SD is the standard deviation of split RMSE, not a standard error based on independent datasets. PCR denotes principal-component ridge regression.

Retained Points	Method	Mean RMSE	SD(RMSE)	Mean MAE
10	PCR	3.315	0.324	2.552
10	Distance taper	9.001	0.860	6.863
10	Functional kernel	10.743	0.802	8.808
10	Rank kNN	9.202	0.820	7.056
10	Uniform kNN	9.683	0.825	7.529
20	PCR	3.008	0.299	2.300
20	Distance taper	8.608	0.846	6.529
20	Functional kernel	10.572	0.799	8.668
20	Rank kNN	8.786	0.839	6.696
20	Uniform kNN	9.247	0.860	7.198
50	PCR	2.987	0.312	2.290
50	Distance taper	8.453	0.879	6.401
50	Functional kernel	10.514	0.797	8.621
50	Rank kNN	8.592	0.824	6.577
50	Uniform kNN	9.068	0.820	7.081
100	PCR	2.983	0.312	2.287
100	Distance taper	8.438	0.879	6.389
100	Functional kernel	10.507	0.797	8.616
100	Rank kNN	8.584	0.821	6.573
100	Uniform kNN	9.055	0.813	7.073

Figure 5 compares predictive RMSE and perturbation bounds across spectral resolutions.

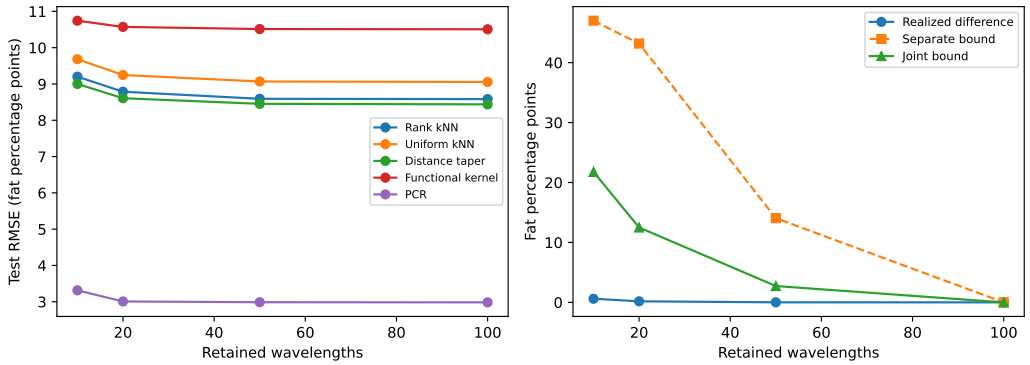


Figure 5. Tecator resolution analysis. **Left**: held-out mean RMSE for five methods. **Right**: mean paired rank-prediction difference (solid line with circular markers) and the certified upper bound (dashed line with square markers), as plotted by the reproducibility script. These are stability certificates, not intervals for latent fat regression.

Because the true conditional mean of fat content is unknown, this application does not estimate coverage of confidence intervals for $m(x)$. Observed fat values are noisy responses, and their inclusion in a mean confidence interval would measure a different target. The repeated splits reuse observations, so their variability is descriptive; no independent-replication significance test is reported.

8.4. Evaluating the Joint Certificate on Real Spectra

We compute both assignment endpoints in Theorem 2 for every held-out spectrum, resolution, and split, keeping the same k and reconstruction-error bound as in the original certificate. This gives 11,600 paired certificate comparisons. No additional outcome-based tuning is introduced; training responses enter the certificate because its target is the perturbation of the fitted prediction on those same responses.

Table 16 presents joint-assignment certificate results.

Table 16. Tecator prediction perturbations and upper bounds, averaged over the same 50 splits and 58 test spectra per split. Units are fat percentage points.

Retained Points	Actual Difference	Separate Bound	Joint Bound
10	0.6038	46.9974	21.7635
20	0.1751	43.1873	12.4981
50	0.0070	14.0498	2.7261
100	0.0000	0.0000	0.0000

The joint bound decreases from 46.997 to 21.764 at $p = 10$, from 43.187 to 12.498 at $p = 20$, and from 14.050 to 2.726 at $p = 50$, relative to the separate bound. These are mean reductions of approximately 54%, 71%, and 81%. Every paired prediction difference is below the joint bound, and every joint bound is below the separate bound up to numerical tolerance 10^{-9}. At full resolution, both bounds are zero up to floating-point rounding. Across the 50 repeated split-specific means, the descriptive standard deviations of the joint bound are 1.362, 1.028, and 0.344 at $p = 10, 20, 50$, respectively; the corresponding standard deviations of the realized mean perturbation are 0.173, 0.086, and 0.012. The 5th–95th percentile ranges of the split-level mean joint bound are 19.110–23.280, 10.591–13.853, and 2.254–3.346. These summarize split-to-split variability and are not independent-dataset confidence intervals.

Averages alone conceal substantial slack. Table 17 reports medians, 95th percentiles, maxima, and the ratio of joint certificate to realized perturbation for cases with nonzero

realized perturbation. At $p = 10$, for example, the median realized perturbation is zero, and its 95th percentile is 3.081, whereas the median and 95th percentile joint bounds are 22.013 and 34.907. The corresponding nonzero-case ratio has a median of 17.9 and a 95th percentile of 214.3. Thus, the certificate must not be interpreted as a predictor of realized perturbation.

Table 17. Distribution of Tecator realized perturbations and joint certificates. Ratios are computed only when the realized perturbation exceeds 10^{-10}; zero realized differences are common and make an ordinary ratio undefined.

p	Act. med.	Act. q95	Act. max	Cert. med.	Cert. q95	Cert. max	Ratio med.	Ratio q95
10	0.000	3.081	12.727	22.013	34.907	41.193	17.9	214.3
20	0.000	1.420	7.387	11.923	24.328	36.393	11.1	107.6
50	0.000	0.000	1.700	2.040	7.414	15.273	6.7	116.4

The practical interpretation is a one-sided tolerance audit of a fixed fitted predictor (Table 18). Before examining the realized perturbation, a user specifies an application-dependent maximum acceptable change T in predicted fat percentage points. If the joint bound is at most T, coarsening is certified not to change this fitted prediction by more than T under the stated reference experiment; if the bound exceeds T, the result is simply inconclusive. At $p = 50$, the joint certificate is at most 2 points for 49.3% of the 2900 held-out cases, at most 3 points for 66.7%, and at most 5 points for 84.9%. At $p = 10$, those proportions are only 0.1%, 0.6%, and 1.5%. This quantifies how often the sufficient certificate meets a specified tolerance. Since the full reference spectra are available in this experiment, the realized difference can also be computed directly. The certificate additionally controls every ordering in its stated uncertainty relaxation. Prospective savings in measurement cost would require an independently validated reconstruction bound; this experiment does not establish such savings. No claim is made that 2, 3, or 5 percentage points are universal domain tolerances; they are illustrative thresholds that a user would replace with a scientifically justified tolerance.

Table 18. Fraction of Tecator held-out cases whose joint certificate meets a specified tolerance. Entries are percentages of 2900 repeated test cases per resolution. Tolerances are illustrative fat-percentage-point thresholds, not domain standards.

Retained Points	$C^{\text{joint}} \leq 2$	$C^{\text{joint}} \leq 3$	$C^{\text{joint}} \leq 5$
10	0.14	0.62	1.52
20	2.55	4.62	14.34
50	49.34	66.72	84.86

8.5. Common and Target-Specific Reference Tolerances

Table 19 recomputes all 11,600 joint certificates using $\delta_j = e_{\text{tr,max}} + e_j$ for the target in question, at the same selected k. This is valid on the controlled reference data by the triangle inequality, and its feasible rank sets are subsets of those obtained from the larger split-wide tolerance. Every target-specific bound contains its realized prediction difference and is no larger than its split-wide counterpart up to numerical tolerance.

At $p = 10, 20, 50$, mean joint bounds decrease from 21.764, 12.498, and 2.726 to 20.183, 9.894, and 2.136. Target-specific allowances remove one source of slack but remain much larger than the mean realized changes of 0.604, 0.175, and 0.007. The remaining gap reflects worst-case ordering relaxation and response variability, not an error in empirical prediction. These computations do not supply prospective calibration when reference spectra are unavailable.

Table 19. Tecator reference tolerance audit over the same repeated held-out cases. Error allowances use RMS absorbance; prediction changes and certificates use fat percentage points. Both tolerances require the available full measured reference spectra and are retrospective.

p	Actual Change	Mean δ_{split}	Mean δ_j	Split Joint Bound	Target Joint Bound
10	0.6038	0.07835	0.05903	21.764	20.183
20	0.1751	0.02166	0.01685	12.498	9.894
50	0.0070	0.00456	0.00351	2.726	2.136
100	0.0000	0.00000	0.00000	0.000	0.000

8.6. Reproducibility Materials

Supplementary File S1 contains the Python 3.12.14 source code, Tecator input files and checksums, grouped-data mapping, fixed split indices, replication-level CSV files, all table and figure generators, and numerical checks. Running `python reproduce.py` executes the complete workflow and regenerates every result used in the manuscript. The final archived results were generated with Python 3.12.14, NumPy 2.3.5, SciPy 1.17.0, pandas 2.2.3, scikit-learn 1.8.0, and Matplotlib 3.10.8. The optional source-data conversion uses the version of `rdata` recorded in `requirements-source.txt`. Seed-construction rules are explicit in the scripts and metadata. The six simulation experiments contain 20,175 generated datasets in total (Table 20); evaluating multiple methods, targets, neighborhood sizes, or calibration factors does not create additional independent datasets. The earlier baseline and refinement experiments used 18,600 datasets; adding 135 target training samples and 360 calibration samples gave 19,095. The present three-mechanism extension adds 1080, giving 20,175. The submission abstract supplied with this revision uses the final total.

The joint certificate solves two linear assignment problems. For each observation i and possible rank j, a position is feasible exactly when $\ell_i \le j \le u_i$. The minimization uses cost $Y_i v_j$, the maximization uses its negative, and infeasible pairs receive infinite cost. Both are solved with `scipy.optimize.linear_sum_assignment`. The package also includes 50 exhaustive small-n permutation checks and 100 numerical comparisons of the threshold optimizer with a generic constrained optimizer. Tecator preprocessing is fully scripted: identical 100-point spectra are grouped before splitting; their fat labels are averaged; retained wavelength indices are approximately equally spaced including endpoints; and linear interpolation returns coarsened spectra to the common 100-point grid before RMS distances are computed.

Table 20. Independent simulated datasets and additional evaluations. Repeated Tecator splits are reference-data audits and are not counted as new simulated datasets.

Experiment	Independent Datasets
Baseline factorial prediction/interval study	10,800
Radius-optimization study	5400
One-coefficient resolution study	2400
Target-location study (three targets per training sample)	135
Calibration sensitivity (12 settings per dataset)	360
Alternative reconstructions (nine cells, 120 per cell)	1080
Total	20,175

The neighborhood sensitivity study reuses 3600 baseline datasets and gives 18,000 evaluations. S1 includes the new derivations, complete scenario-specific comparisons, paired MSE standard errors, all neighborhood sensitivity records, the exact-curve coverage comparator, the 108 joint stress-test cases, and 11,600 target-specific Tecator certificates. Its code, tables, figures, and software metadata match this revised manuscript.

With 135 training spectra, the median time per target for rank limits, dense matrix construction, and both assignment solves was 0.90 ms (95th percentile 2.36 ms) across the 8700 non-full-resolution cases in the target-specific audit. Distances and cross-validation are excluded. Timings are specific to the supplied execution environment and single-thread configuration. A representative $p = 20$ call had a Python-tracked allocation peak of 456 KiB; this is not total process memory. A single 135×135 float64 cost matrix uses 145,800 bytes (about 142 KiB) and the Boolean feasibility mask 18,225 bytes (about 18 KiB). Temporary masked costs and solver work arrays add to these allocations. Each certificate is solved independently; the implementation does not retain dense assignment matrices for all targets at once. Root tolerances and direct evaluation of $A(w)$ are documented in Section 6.1 and in `code/certified_optimization.py`.

9. Discussion

The main practical distinction is between localized estimation, same-sample stability, and calibrated uncertainty. A reconstructed estimator can have a direct Gaussian limit even when its exact-curve counterpart is not asymptotically identical on the same sample. This follows because selected observations may still approach the target and the squared weights remain fixed. In contrast, paired equivalence requires control of weight differences and therefore of rank perturbations.

Certified intervals provide a different route: their finite-sample mean coverage requires valid smoothness, conditional noise, and reconstruction-error bounds independently of the response errors used in the interval. They explicitly add a bias allowance derived from smoothness and a reconstruction-error certificate. This protects coverage without an undersmoothing approximation, but requires credible constants and can produce wide intervals. The Monte Carlo results illustrate both facts; coverage equal to one reflects wide worst-case allowances rather than verified practical calibration. The irregular-grid, Laplace-noise, and truncated-representation extensions show that available error bounds can preserve the guarantee while making intervals still less informative. The joint assignment certificate and radius-optimized weights reduce two explicit sources of conservatism. Independent calibration can provide a valid upper bound for the reconstruction noise scale in the finite-rank Gaussian example, but general data-driven certification of the global smoothness constant and a uniform conditional sub-Gaussian response bound remains unresolved. Without such external calibration, Theorems 8 and 9 are best viewed as benchmark guarantees and sensitivity tools rather than fully automatic procedures.

The contribution combines local rank control with a quantitative resolution condition and joint permutation certificates. Existing verification work [13] provides relevant background; the comparison concerns the stated uncertainty sets and guarantees, not priority for kNN certification in general. Its confidence-radius component adapts the established threshold optimization to reconstruction-adjusted bias bounds. The estimator belongs to a pre-existing weighted-neighbor family, and transfer from reconstructed curves has prior literature. We do not establish minimax optimality, an adaptive oracle risk inequality, robustness to arbitrary response contamination, or simultaneous coverage over all functional targets. The real-data benchmark concerns controlled coarsening of measured spectra and does not identify the latent continuous measurement error.

10. Conclusions

The contribution quantifies reconstruction effects on an existing rank-weighted functional smoother rather than proposing a new estimator. Deterministic rank limits and joint assignment bound same-sample fitted-value changes; conditional localization gives distinct sufficient conditions for direct inference and paired equivalence. The explicit local rate

requires polynomial small balls and a radial-density bound and is not a general infinite-dimensional rate. The threshold optimizer is an adaptation of an established weighting principle. Its intervals provide finite-sample coverage for $m(x)$ only under independently valid smoothness, conditional response-noise, and reconstruction-error bounds, with calibration failure included in the budget. The simulations show substantial conservatism even after optimization and under alternative reconstruction mechanisms. Predictive rankings vary by target; PCR is stronger on Tecator. That application is a controlled coarsening stability audit of measured spectra, not validation of latent continuous-spectrum inference or prospective reconstruction calibration. These limitations define where the results can support an audit and where further calibration and methodological work are needed.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/math14213969/s1, Supplementary File S1 (reproducibility archive): Python scripts for simulation, estimator evaluation, interval optimization, figure generation, and validation; Tecator data files (original and deduplicated spectra), documentation, and checksum information; simulation and Tecator replication-level CSV files; recorded split indices and numerical result summaries; and computational metadata and validation reports. The five supplementary figure files are: `simulation_rmse.pdf` (RMSE by scenario and sample size), `stability_coverage.pdf` (prediction perturbation and nominal interval coverage), `optimized_width.pdf` (optimized relative interval width), `tecator_data.pdf` (measured spectra and observed-versus-predicted fat), and `tecator_performance.pdf` (test RMSE and perturbation certificates). These figures are also reproduced in the article.

Author Contributions: Conceptualization, W.B.; methodology, W.B.; software, W.B.; validation, F.A.; formal analysis, F.A.; investigation, F.A.; resources, F.A.; data curation, W.B.; writing—original draft preparation, F.A.; writing—review and editing, W.B.; visualization, F.A.; supervision, W.B. and F.A.; project administration, F.A.; funding acquisition, F.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R358), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Data Availability Statement: The Tecator data used in this study are publicly available through the CRAN `fda.usc` distribution identified in Reference [20,21]. The exact source files, checksums, preprocessing outputs, simulation seeds, replication-level results, and Python code used to reproduce the analyses are provided in Supplementary File S1.

Acknowledgments: The authors extend their sincere appreciation to Princess Nourah bint Abdulrahman University Researchers Supporting Project number PNURSP2026R358, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Laloë, T. A k-nearest neighbor approach for functional regression. *Stat. Probab. Lett.* **2008**, *78*, 1189–1193. [CrossRef]
2. Burba, F.; Ferraty, F.; Vieu, P. k-Nearest Neighbour method in functional nonparametric regression. *J. Nonparametr. Stat.* **2009**, *21*, 453–469. [CrossRef]
3. Biau, G.; Cérou, F.; Guyader, A. Rates of convergence of the functional k-nearest neighbor estimate. *IEEE Trans. Inf. Theory* **2010**, *56*, 2034–2040. [CrossRef]
4. Kara-Zaitri, L.; Laksaci, A.; Rachdi, M.; Vieu, P. Data-driven kNN estimation in nonparametric functional data analysis. *J. Multivar. Anal.* **2017**, *153*, 176–188. [CrossRef]
5. Kudraszow, N.L.; Vieu, P. Uniform consistency of kNN regressors for functional variables. *Stat. Probab. Lett.* **2013**, *83*, 1863–1870. [CrossRef]
6. Ferraty, F.; Vieu, P. *Nonparametric Functional Data Analysis: Theory and Practice*; Springer: New York, NY, USA, 2006. [CrossRef]

7. Ferraty, F.; Laksaci, A.; Tadj, A.; Vieu, P. Rate of uniform consistency for nonparametric estimates with functional variables. *J. Stat. Plan. Inference* **2010**, *140*, 335–352. [CrossRef]
8. Carroll, R.J.; Härdle, W. Symmetrized nearest neighbor regression estimates. *Stat. Probab. Lett.* **1989**, *7*, 315–318. [CrossRef]
9. Mas, A. Lower bound in regression for functional data by representation of small ball probabilities. *Electron. J. Stat.* **2012**, *6*, 1745–1778. [CrossRef]
10. Meister, A. Optimal classification and nonparametric regression for functional data. *Bernoulli* **2016**, *22*, 1729–1744. [CrossRef]
11. Forzani, L.; Fraiman, R.; Llop, P. Nonparametric regression based on discretely sampled curves. *REVSTAT Stat. J.* **2020**, *18*, 1–26. [CrossRef]
12. Xing, Y.; Song, Q.; Cheng, G. Predictive power of nearest neighbors algorithm under random perturbation. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*; Proceedings of Machine Learning Research; PMLR: Cambridge, MA, USA, 2021; Volume 130, pp. 496–504. Available online: https://proceedings.mlr.press/v130/xing21a.html (accessed on 25 October 2026).
13. Fassina, N.; Ranzato, F.; Zanella, M. Robustness verification of k-nearest neighbors by abstract interpretation. *Knowl. Inf. Syst.* **2024**, *66*, 4825–4859. [CrossRef]
14. Armstrong, T.B.; Kolesár, M. Optimal inference in a class of regression models. *Econometrica* **2018**, *86*, 655–683. [CrossRef]
15. Schennach, S.M. A bias bound approach to non-parametric inference. *Rev. Econ. Stud.* **2020**, *87*, 2439–2472. [CrossRef]
16. Anava, O.; Levy, K.Y. k*-Nearest Neighbors: From Global to Local. In *Advances in Neural Information Processing Systems*; NIPS Foundation: San Diego, CA, USA, 2016; Volume 29, pp. 4923–4931. Available online: https://papers.nips.cc/paper_files/paper/2016/hash/2c6ae45a3e88aee548c0714fad7f8269-Abstract.html (accessed on 25 October 2026).
17. Elmakias, I.; Vilenchik, D. Choosing the right dataset: Hardness criteria for feature selection benchmarking. *Knowl.-Based Syst.* **2026**, *334*, 115022. [CrossRef]
18. Vilenchik, D.; Yichye, B.; Abutbul, M. To Interpret or Not to Interpret PCA? This Is Our Question. *Proc. Int. AAAI Conf. Web Soc. Media* **2019**, *13*, 655–658. [CrossRef]
19. Febrero-Bande, M.; Oviedo de la Fuente, M. Statistical computing in functional data analysis: The R package fda.usc. *J. Stat. Softw.* **2012**, *51*, 1–28. [CrossRef]
20. CRAN fda.usc Mirror. Tecator Data and Documentation. Data File. Available online: https://github.com/cran/fda.usc/blob/master/data/tecator.rda (accessed on 17 September 2026).
21. CRAN fda.usc Mirror. Tecator Data and Documentation. Documentation. Available online: https://github.com/cran/fda.usc/blob/master/man/tecator.Rd (accessed on 17 September 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.