fractal and fractional

MDPI

Article

Fractal Perturbation of the Nadaraya–Watson Estimator

Dah-Chin Luor [1,*,†] and Chiao-Wen Liu [2,†]

[1] Department of Data Science and Analytics, School of Intelligent Science and Technology, I-Shou University, Dashu District, Kaohsiung City 84001, Taiwan
[2] Department of Applied Science, School of Academic Studies, R.O.C. Naval Academy, Zuoying District, Kaohsiung City 813000, Taiwan
* Correspondence: dclour@isu.edu.tw
† These authors contributed equally to this work.

Abstract: One of the main tasks in the problems of machine learning and curve fitting is to develop suitable models for given data sets. It requires to generate a function to approximate the data arising from some unknown function. The class of kernel regression estimators is one of main types of nonparametric curve estimations. On the other hand, fractal theory provides new technologies for making complicated irregular curves in many practical problems. In this paper, we are going to investigate fractal curve-fitting problems with the help of kernel regression estimators. For a given data set that arises from an unknown function m, one of the well-known kernel regression estimators, the Nadaraya–Watson estimator $\hat{m}$, is applied. We consider the case that m is Hölder-continuous of exponent β with $0 < \beta \leq 1$, and the graph of m is irregular. An estimation for the expectation of $|\hat{m} - m|^2$ is established. Then a fractal perturbation $f_{[\hat{m}]}$ corresponding to $\hat{m}$ is constructed to fit the given data. The expectations of $|f_{[\hat{m}]} - \hat{m}|^2$ and $|f_{[\hat{m}]} - m|^2$ are also estimated.

Keywords: kernel regression estimators; Nadaraya–Watson estimator; fractal interpolation; curve fitting

check for updates

Citation: Luor, D.-C.; Liu, C.-W. Fractal Perturbation of the Nadaraya–Watson Estimator. *Fractal Fract.* **2022**, *6*, 680. https://doi.org/10.3390/fractalfract6110680

Academic Editor: Viorel-Puiu Paun

Received: 8 October 2022
Accepted: 15 November 2022
Published: 17 November 2022

1. Introduction

One of the main tasks in the problems of machine learning, curve fitting, signal analysis, and many statistical applications is to develop suitable models for given data sets. In many real-world applications, it requires to generate a function to interpolate or to approximate the data arising from some unknown function. In data-fitting problems, interpolation is usually applied when the data are noise-free, and regression is considered if we have noisy observations.

The theory of nonparametric modeling of a regression has been developed by many researchers. Several types of estimators and their statistical properties have been studied in the literature. The class of kernel estimators is one of the main types of nonparametric curve estimations, and the Nadaraya–Watson estimator, the Priestley–Chao estimator, and the Gasser–Müller estimator are widely used in applications. See [1–6] and references given in these books. In [7,8], the authors investigated the differences between several types of kernel regression estimators, and there is no answer to which of these estimators is the best since each of them has advantages and disadvantages.

Fractal theory provides another technology for making complicated curves and fitting experimental data. A fractal interpolation function (FIF) is a continuous function interpolating a given set of points, and the graph of a FIF is the attractor of an iterated function system. The concept of FIFs was introduced by Barnsley ([9,10]), and it has been developed to be the basis of an approximation theory for nondifferentiable functions. FIFs can also be applied to model discrete sequences ([11–13]). Various types of FIFs and their approximation properties were discussed in [14–44], and the references given in the literature. See also the book [45] for recent developments. In [46–50], the construction of FIFs for random data sets is given, and some statistical properties of such FIFs were investigated. In [51], the authors made a topological–geometric contribution for the development and applications of fractal models, which present periodic changes.

For a given data set that arises from an unknown function m, the purpose of this paper is not to establish a fractal function that interpolates points in the data set, but we aim to find a fractal function that has good approximation for these data points. In [52], the authors trained SVM by the chosen training data and then applied the SVM model to calculate the interpolation points used to construct a linear FIF. In this paper, we consider the Nadaraya–Watson estimator $\hat{m}$ for some sample data chosen from a given data set, and establish an estimation for the expectation of $|\hat{m} - m|^2$. Then a FIF $f_{[\hat{m}]}$ corresponding to $\hat{m}$ is constructed to fit the given data set, and the expectations of $|f_{[\hat{m}]} - \hat{m}|^2$ and $|f_{[\hat{m}]} - m|^2$ are also estimated.

Throughout this paper, let $\mathcal{D} = \{(t_i, y_i) \in \mathbb{R} \times \mathbb{R} : i = 0, 1, \cdots, N\}$ be a given data set, where N is an integer greater than or equal to 2, and $t_0 < t_1 < \cdots < t_N$. We take $t_0 = 0$ and $t_N = 1$ for convenience. Let $I = [0, 1]$ and $I_i = [t_{i-1}, t_i]$ for $i = 1, \cdots, N$. Let $C[I]$ denote the set of all real-valued continuous functions defined on I. The set of functions in $C[I]$ that interpolate all points in $\mathcal{D}$ is denoted by $C_{\mathcal{D}}[I]$. Define $\|f\|_\infty = \max_{t \in I} |f(t)|$ for $f \in C[I]$. It is known that $(C[I], \|\cdot\|_\infty)$ is a Banach space, and $C_{\mathcal{D}}[I]$ is a complete metric space, where the metric is induced by $\|\cdot\|_\infty$.

2. Construction of Fractal Interpolation Functions

In this section, we establish a fractal perturbation of a given function in $C[I]$. The construction given here has been treated in the literature (see [47]). We show the details here to make our paper more self-contained.

Let $u \in C[I]$ and $\mathcal{D} = \{(t_i, y_i) : y_i = u(t_i), i = 0, 1, \cdots, N\}$, where $0 = t_0 < t_1 < \cdots < t_N = 1$. Assume that the data points in $\mathcal{D}$ are non-collinear. For $i = 1, \cdots, N$, let $L_i : I \to I_i$ be a homeomorphism such that $L_i(0) = t_{i-1}$ and $L_i(1) = t_i$. Define $M_i : I \times \mathbb{R} \to \mathbb{R}$ by

$$M_i(t, y) = s_i y + u(L_i(t)) - s_i p(t), \tag{1}$$

where $-1 < s_i < 1$ and p is a continuous function on I such that $p(0) = u(0)$ and $p(1) = u(1)$. Then $M_i(0, u(0)) = y_{i-1}$, $M_i(1, u(1)) = y_i$, and

$$|M_i(t, y) - M_i(t, y^*)| = |s_i||y - y^*| \text{ for all } t \in I \text{ and } y, y^* \in \mathbb{R}. \tag{2}$$

Define $W_i : I \times \mathbb{R} \to I_i \times \mathbb{R}$ by $W_i(t, y) = (L_i(t), M_i(t, y))$ for $i = 1, \cdots, N$. For $h \in C_{\mathcal{D}}[I]$, let $G_h = \{(t, h(t)) : t \in I\}$. Then $W_i(G_h) = \{(L_i(t), M_i(t, h(t))) : t \in I\}$. Since $L_i : I \to I_i$ is a homeomorphism, $W_i(G_h)$ can be written as

$$W_i(G_h) = \{(t, M_i(L_i^{-1}(t), h(L_i^{-1}(t)))) : t \in I_i\}.$$

Hence $W_i(G_h)$ is the graph of the continuous function $h_i : I_i \to \mathbb{R}$ defined by $h_i(t) = M_i(L_i^{-1}(t), h(L_i^{-1}(t)))$. Define a mapping $T : C_{\mathcal{D}}[I] \to C_{\mathcal{D}}[I]$ by

$$T(h)(t) = h_i(t) = s_i h(L_i^{-1}(t)) + u(t) - s_i p(L_i^{-1}(t)),\ t \in I_i. \tag{3}$$

By (3) we see that, for $g, h \in C_{\mathcal{D}}[I]$ and $t \in I_i$,

$$|T(g)(t) - T(h)(t)| \le |s_i||g(L_i^{-1}(t)) - h(L_i^{-1}(t))|.$$

Then

$$\|T(g) - T(h)\|_\infty \le \max_{i=1,\cdots,N} |s_i| \left\{ \max_{z \in I} |g(z) - h(z)| \right\} \le s\|g - h\|_\infty.$$

Here $s = \max\{|s_1|, \cdots, |s_N|\}$. Since $0 \le s < 1$, we have the following theorem ([47], Theorem 2.1).

Theorem 1. *The operator T given by* (3) *is a contraction mapping on $C_{\mathcal{D}}[I]$.*

Definition 1. *The fixed point $f_{[u]}$ of T in $C_{\mathcal{D}}[I]$ is called a fractal interpolation function (FIF) on I corresponding to the continuous function u.*

The FIF $f_{[u]}$ given in Definition 1 satisfies the following equation for $i = 1, \cdots, N$:

$$f_{[u]}(t) = s_i \left\{ f_{[u]}(L_i^{-1}(t)) - p(L_i^{-1}(t)) \right\} + u(t), \quad t \in I_i. \tag{4}$$

If $s_i = 0$ for all i, then $f_{[u]} = u$. Therefore, $f_{[u]}$ can be treated as a fractal perturbation of u.

3. The Nadaraya–Watson Estimator

Let $\mathcal{D} = \{(t_i, y_i) \in \mathbb{R} \times \mathbb{R} : i = 0, 1, \cdots, N\}$ be a given data set, where $0 = t_0 < t_1 < \cdots < t_N = 1$. Suppose that

$$Y_i = m(t_i) + \epsilon_i, \quad \text{for } i = 0, 1, \cdots, N, \tag{5}$$

where $m : [0,1] \to \mathbb{R}$ is an unknown function, and each y_i is an observation of Y_i. Here, all ϵ_i are independent stochastic disturbance terms with zero expectation, $\mathbf{E}[\epsilon_i] = 0$, and finite variance, $\mathbf{Var}[\epsilon_i] \leq \sigma^2 < \infty$. In this section, we consider the Nadaraya–Watson estimator $\hat{m}$ for $\mathcal{D}$ and establish an estimation for the expectation of $|\hat{m} - m|^2$.

Consider the case that m is Hölder continuous of exponent β with $0 < \beta \leq 1$, and the graph of m is irregular. Then, m satisfies the inequality with $0 < \beta \leq 1$ and $\lambda > 0$:

$$|m(t) - m(t')| \leq \lambda |t - t'|^{\beta}, \quad t, t' \in I. \tag{6}$$

The Nadaraya–Watson estimator $\hat{m}$ of m is defined by

$$\hat{m}(t) = \frac{\sum_{i=0}^{N} k_d(t - t_i) Y_i}{\sum_{j=0}^{N} k_d(t - t_j)}, \quad \text{where } k_d(z) = \frac{1}{d} k\left(\frac{z}{d}\right). \tag{7}$$

Here $d > 0$ is a bandwidth, and k is an integrable function defined on $\mathbb{R}$.

The function k is called a kernel and is usually assumed to be bounded and satisfies some integrable conditions. Some widely used kernels are given in ([2], p. 41) and ([5], p. 3), and the estimations using different kernels are usually numerically similar (see [6]). In this paper, we assume that there are positive numbers C_1, C_2, η, and R such that the kernel k satisfies the condition

$$C_1 \chi_{[-\eta, \eta]}(z) \leq k(z) \leq C_2 \chi_{[-R, R]}(z), \quad z \in \mathbb{R}. \tag{8}$$

Condition (8) and its multidimensional form was considered in ([5], Theorem 1.7) and ([1], Theorem 5.1).

A new estimation for the bias of $\hat{m}$ was obtained in [53]. Here, we give an estimation for $\mathbf{E}[(\hat{m}(t) - m(t))^2]$ in the following Theorem 2. Similar results were studied in [1,2,5], and other literature. The convergence rate of upper estimation obtained in Theorem 2 is the same as the known results.

The Nadaraya–Watson estimator $\hat{m}$ given in (7) can be written in the form

$$\hat{m}(t) = \sum_{i=0}^{N} W_i(t) Y_i, \quad \text{where} \quad W_i(t) = \frac{k_d(t - t_i)}{\sum_{j=0}^{N} k_d(t - t_j)}. \tag{9}$$

Then $\sum_{i=0}^{N} W_i(t) = 1$ for all t and

$$\mathbf{E}[\hat{m}(t)] = \sum_{i=0}^{N} W_i(t) \mathbf{E}[Y_i] = \sum_{i=0}^{N} W_i(t) m(t_i). \tag{10}$$

In the following lemma, we give a lower bound for $\sum_{j=0}^{N} k_d(t-t_j)$. Define

$$a_N = \min_{1\le k\le N} t_k - t_{k-1}, \quad A_N = \max_{1\le k\le N} t_k - t_{k-1}. \tag{11}$$

Lemma 1. *Let $0 = t_0 < t_1 < \cdots < t_N = 1$. Suppose that $k : \mathbb{R} \to \mathbb{R}$ and there are positive numbers C_1 and η such that $C_1\chi_{[-\eta,\eta]}(z) \le k(z)$ for $z \in \mathbb{R}$. Let $d > 0$ and let A_N and k_d be defined in (11) and (7), respectively. Assume that $A_N < 2d\eta$ and $A_N \le \frac{\alpha}{N}$ for some $\alpha > 0$. Then for $0 \le t \le 1$,*

$$\sum_{j=0}^{N} k_d(t-t_j) \ge \frac{C_1\eta N}{\alpha}. \tag{12}$$

Proof. For $0 \le t \le 1$, the condition $C_1\chi_{[-\eta,\eta]}(z) \le k(z)$ implies that

$$\sum_{j=0}^{N} k_d(t-t_j) = \frac{1}{d}\sum_{j=0}^{N} k\left(\frac{t-t_j}{d}\right) \ge \frac{C_1}{d}\sum_{j=0}^{N}\chi_{[-\eta,\eta]}\left(\frac{t-t_j}{d}\right) = \frac{C_1}{d}|E_\eta(t)|,$$

where $E_\eta(t) = \{t_j : j = 0, 1, \cdots, N, \text{ and } |\frac{t-t_j}{d}| \le \eta\}$ and $|E_\eta(t)|$ is the number of elements of $E_\eta(t)$. Since $|\frac{t-t_j}{d}| \le \eta$ if and only if $t_j \in [t-d\eta, t+d\eta] \cap [0,1]$, we have $E_\eta(t) = \{t_j : j = 0, 1, \cdots, N, \text{ and } t_j \in [t-d\eta, t+d\eta] \cap [0,1]\}$.

For $t \in [d\eta, 1-d\eta]$, we have $[t-d\eta, t+d\eta] \subseteq [0,1]$, and by the condition $A_N < 2d\eta$, we see that $|E_\eta(t)| \ge [\frac{2d\eta}{A_N}] \ge 1$ and this implies $|E_\eta(t)| \ge \frac{d\eta}{A_N}$. For $t \in [0, d\eta)$, we have $[t-d\eta, t+d\eta] \cap [0,1] = [0, t+d\eta]$ and $t_0 = 0 \in E_\eta(t)$. Hence $|E_\eta(t)| \ge [\frac{t+d\eta}{A_N}] + 1 \ge 1$ and $|E_\eta(t)| \ge \frac{d\eta}{A_N}$. For $t \in (1-d\eta, 1]$, we have $[t-d\eta, t+d\eta] \cap [0,1] = [t-d\eta, 1]$ and $t_N = 1 \in E_\eta(t)$. Hence $|E_\eta(t)| \ge [\frac{1-t+d\eta}{A_N}] + 1 \ge 1$ and $|E_\eta(t)| \ge \frac{d\eta}{A_N}$. Then the condition $A_N \le \frac{\alpha}{N}$ implies (12). □

Theorem 2. *Let $\mathcal{D}$ be a given data set and assume that m satisfies (6). Suppose that k satisfies (8) and $\hat{m}$ is defined by (7). Assume that $A_N < 2d\eta$ and $A_N \le \frac{\alpha}{N}$ for some $\alpha > 0$. Then we have*

$$\mathbf{E}[(\hat{m}(t) - m(t))^2] \le \lambda^2 R^{2\beta} d^{2\beta} + \left(\frac{\alpha C_2\sigma^2}{C_1\eta}\right)\frac{1}{Nd}. \tag{13}$$

Proof. We see that

$$\mathbf{E}[(\hat{m}(t) - m(t))^2] = \{\mathbf{E}[\hat{m}(t)] - m(t)\}^2 + \mathbf{E}[\hat{m}(t)^2] - (\mathbf{E}[\hat{m}(t)])^2. \tag{14}$$

By (6) and (9)–(10), we have

$$|\mathbf{E}[\hat{m}(t)] - m(t)| = \left|\sum_{i=0}^{N} W_i(t)(m(t_i) - m(t))\right| \le \lambda\sum_{i=0}^{N} W_i(t)|t_i - t|^\beta.$$

Condition (8) implies that $k(\frac{t-t_i}{d}) = 0$ if $|\frac{t-t_i}{d}| > R$. Therefore,

$$|\mathbf{E}[\hat{m}(t)] - m(t)| \le \lambda d^\beta \frac{\sum_{i=0}^{N} k(\frac{t-t_i}{d})|\frac{t-t_i}{d}|^\beta}{\sum_{j=0}^{N} k(\frac{t-t_j}{d})} \le \lambda R^\beta d^\beta. \tag{15}$$

On the other hand, by (8) and (12), we also have

$$\sup_{i,t} W_i(t) = \sup_{i,t}\frac{k(\frac{t-t_i}{d})}{\sum_{j=0}^{N} k(\frac{t-t_j}{d})} \le \frac{\alpha C_2}{C_1\eta N d}. \tag{16}$$

By (9), (10) and (5), we have

$$\mathbf{E}[\hat{m}(t)^2] - (\mathbf{E}[\hat{m}(t)])^2 = \mathbf{E}[(\hat{m}(t) - \mathbf{E}[\hat{m}(t)])^2] = \mathbf{E}\left[\left(\sum_{i=0}^{N} W_i(t)\epsilon_i\right)^2\right].$$

Since all ϵ_i are independent and satisfy $\mathbf{E}[\epsilon_i] = 0$ and $\mathbf{Var}[\epsilon_i] \leq \sigma^2 < \infty$, the condition $\sum_{i=0}^{N} W_i(t) = 1$ and estimation (16) imply that

$$\mathbf{E}[\hat{m}(t)^2] - (\mathbf{E}[\hat{m}(t)])^2 = \sum_{i=0}^{N} W_i(t)^2 \mathbf{E}[\epsilon_i^2] \leq \sigma^2 \left(\sup_{i,t} W_i(t)\right) \sum_{i=0}^{N} W_i(t) \leq \left(\frac{\alpha C_2 \sigma^2}{C_1 \eta}\right) \frac{1}{Nd}.$$

Then by (14) and (15), we have (13). □

For a given kernel k which satisfies (8), estimation (13) shows that C_1 and η should be chosen so that $C_1\eta$ is as large as possible. The minimizer d^* with respect to d of the right-hand side of (13) can be obtained by setting $E(d) = \lambda^2 R^{2\beta} d^{2\beta} + (\frac{\alpha C_2 \sigma^2}{C_1 \eta N}) d^{-1}$, and then solve the equation

$$E'(d) = (2\beta)\lambda^2 R^{2\beta} d^{2\beta - 1} - \left(\frac{\alpha C_2 \sigma^2}{C_1 \eta N}\right) d^{-2} = 0.$$

We have

$$d^* = \left(\frac{\alpha C_2 \sigma^2}{2\beta C_1 \eta \lambda^2 R^{2\beta}}\right)^{\frac{1}{2\beta+1}} N^{\frac{-1}{2\beta+1}} \tag{17}$$

and the upper estimate given in (13) can be reduced to $C^* N^{-2\beta/(2\beta+1)}$, where C^* depends on α, β, λ, σ^2, η, R, C_1, and C_2.

4. Fractal Perturbation of the Nadaraya–Watson Estimator

In this section, we consider FIFs $f_{[\hat{m}]}$ corresponding to the function $\hat{m}$ and we establish estimations for the expectation of $|f_{[\hat{m}]} - \hat{m}|^2$ and $|f_{[\hat{m}]} - m|^2$. Suppose that k is continuous and we replace each Y_i in (7) by y_i. Then $\hat{m} \in C[I]$. By the construction given in Section 2 with $u = \hat{m}$, we have a FIF $f_{[\hat{m}]}$ on I that satisfies the equation for $i = 1, \cdots, N$:

$$f_{[\hat{m}]}(t) = s_i \left\{ f_{[\hat{m}]}(L_i^{-1}(t)) - p(L_i^{-1}(t)) \right\} + \hat{m}(t), \quad t \in I_i. \tag{18}$$

Here, p is chosen to be the linear polynomial such that $p(0) = \hat{m}(0)$ and $p(1) = \hat{m}(1)$. Then we replace y_i by Y_i for each i and consider $f_{[\hat{m}]}(t)$ a random variable for every $t \in I$. We are interested in estimations for $\|\mathbf{E}[|f_{[\hat{m}]} - m|^2]\|_\infty$.

Theorem 3. *Suppose that k is continuous and k satisfies (8) with $R = 1$ and $C_2 = 1$. Suppose that m satisfies (6) and $\hat{m}$ is defined by (7). Let $M = \max\{|m(t_i)| : i = 0, 1, \cdots, N\}$. Assume that $A_N < 2d\eta$, $A_N \leq \frac{\alpha}{N}$, and $a_N \geq \frac{\tau}{N}$ for some $\alpha > 0$ and $\tau > 0$, where A_N and a_N are defined in (11). Suppose that $0 < s = \max\{|s_1|, \cdots, |s_N|\} < 2^{-1/2}$ and $\|\mathbf{E}[|f_{[\hat{m}]} - \hat{m}|^2]\|_\infty < \infty$. Then we have*

$$\|\mathbf{E}[|f_{[\hat{m}]} - \hat{m}|^2]\|_\infty \leq \left(\frac{72 s^2 \alpha^2 (M^2 + \sigma^2)}{(1 - 2s^2) C_1^2 \eta^2 \tau^2}\right) \frac{(Nd + \tau)^2}{(Nd)^2}, \tag{19}$$

$$\|\mathbf{E}[|f_{[\hat{m}]} - m|^2]\|_\infty \leq \left(\frac{144 s^2 \alpha^2 (M^2 + \sigma^2)}{(1 - 2s^2) C_1^2 \eta^2 \tau^2}\right) \frac{(Nd + \tau)^2}{(Nd)^2} + 2\lambda^2 d^{2\beta} + \left(\frac{2\alpha\sigma^2}{C_1 \eta}\right) \frac{1}{Nd}. \tag{20}$$

Proof. For $t \in I_i$, (18) implies

$$|f_{[\hat{m}]}(t) - \hat{m}(t)|^2 \leq 2s_i^2 \left\{ |f_{[\hat{m}]}(L_i^{-1}(t)) - \hat{m}(L_i^{-1}(t))|^2 + |\hat{m}(L_i^{-1}(t)) - p(L_i^{-1}(t))|^2 \right\},$$

and we have

$$\sup_{t \in I_i} \mathbf{E}[|f_{[\hat{m}]}(t) - \hat{m}(t)|^2] \leq 2s_i^2 \left(\sup_{z \in I} \mathbf{E}[|f_{[\hat{m}]}(z) - \hat{m}(z)|^2] + \sup_{z \in I} \mathbf{E}[|\hat{m}(z) - p(z)|^2] \right).$$

Then

$$\|\mathbf{E}[|f_{[\hat{m}]} - \hat{m}|^2]\|_\infty \leq 2s^2 \{ \|\mathbf{E}[|f_{[\hat{m}]} - \hat{m}|^2]\|_\infty + \|\mathbf{E}[|\hat{m} - p|^2]\|_\infty \}$$

and therefore

$$\|\mathbf{E}[|f_{[\hat{m}]} - \hat{m}|^2]\|_\infty \leq \frac{2s^2}{1 - 2s^2} \|\mathbf{E}[|\hat{m} - p|^2]\|_\infty. \tag{21}$$

Since p is the linear polynomial with $p(0) = \hat{m}(0)$ and $p(1) = \hat{m}(1)$, we have

$$p(t) = \hat{m}(0) + (\hat{m}(1) - \hat{m}(0))t, \quad t \in I, \tag{22}$$

and then

$$|\hat{m}(t) - p(t)| = |(\hat{m}(t) - \hat{m}(0))(1 - t) + (\hat{m}(t) - \hat{m}(1))t|.$$

The convexity of the square function $x \mapsto x^2$ implies that

$$|\hat{m}(t) - p(t)|^2 \leq (1 - t)|\hat{m}(t) - \hat{m}(0)|^2 + t|\hat{m}(t) - \hat{m}(1)|^2$$

and therefore

$$\mathbf{E}[|\hat{m}(t) - p(t)|^2] \leq (1 - t)\mathbf{E}[|\hat{m}(t) - \hat{m}(0)|^2] + t\mathbf{E}[|\hat{m}(t) - \hat{m}(1)|^2], \ t \in I. \tag{23}$$

By (9), $\hat{m}(t) - \hat{m}(1) = \sum_{r=0}^{N} (W_r(t) - W_r(1))Y_r$. By (8) with $R = 1$, we see that if $t_r < 1 - d$, then $\frac{1-t_r}{d} > 1$ and $k(\frac{1-t_r}{d}) = 0$. This implies $W_r(1) = 0$. For $t \in I$, if $t_r \notin [t - d, t + d]$, then $|\frac{t-t_r}{d}| > 1$ and $k(\frac{t-t_r}{d}) = 0$. This implies $W_r(t) = 0$. Then

$$\hat{m}(t) - \hat{m}(1) = \sum_{r \in B_t} (W_r(t) - W_r(1))Y_r, \tag{24}$$

where $B_t = \{r : t_r \in [t - d, t + d] \text{ or } t_r \in [1 - d, 1]\}$. Let $\xi = [\frac{d}{a_N}]$. Then the number of elements in B_t is less than $3(\xi + 1)$.

By (12) and (8) with $C_2 = 1$, we have

$$|W_r(t) - W_r(1)| \leq \frac{k_d(t - t_r)}{\sum_{j=0}^{N} k_d(t - t_j)} + \frac{k_d(1 - t_r)}{\sum_{j=0}^{N} k_d(1 - t_j)} \leq \frac{2\alpha}{C_1 \eta N d}. \tag{25}$$

By (5) we also have $\mathbf{E}[Y_r^2] = m(t_r)^2 + \sigma^2$ for $r = 0, 1, \cdots, N$. Condition (6) shows that m is continuous and therefore m is bounded on I. Then for $t \in I$,

$$\begin{aligned} \mathbf{E}[|\hat{m}(t) - \hat{m}(1)|^2] &\leq \left\{ \sum_{r \in B_t} (W_r(t) - W_r(1))^2 \right\} \left\{ \sum_{r \in B_t} \mathbf{E}[Y_r^2] \right\} \\ &\leq \left(\frac{2\alpha}{C_1 \eta N d} \right)^2 (M^2 + \sigma^2)(3\xi + 3)^2. \end{aligned}$$

We also have the same estimate for $\mathbf{E}[|\hat{m}(t) - \hat{m}(0)|^2]$.

By the condition $a_N \geq \frac{\tau}{N}$, we have $\xi \leq \frac{d}{a_N} \leq \frac{Nd}{\tau}$, and then (23) can be reduced to

$$\mathbf{E}[|\hat{m}(t) - p(t)|^2] \leq \left(\frac{36\alpha^2(M^2+\sigma^2)}{C_1^2\eta^2\tau^2}\right)\frac{(Nd+\tau)^2}{(Nd)^2}, \quad t \in I. \tag{26}$$

Thus, (19) can be obtained by (21) and (26). Moreover, we have (20) by (13), (19), and the inequality

$$\|\mathbf{E}[|f_{[\hat{m}]} - m|^2]\|_\infty \leq 2\|\mathbf{E}[|f_{[\hat{m}]} - \hat{m}|^2]\|_\infty + 2\|\mathbf{E}[|\hat{m} - m|^2]\|_\infty.$$

□

For a given kernel k, which satisfies condition (8), estimation (20) shows that C_1 and η should be chosen so that $C_1\eta$ is as large as possible. If we choose $d = d^*$, where d^* is given by (17) with $C_2 = 1$ and $R = 1$, then (20) can be reduced to

$$\|\mathbf{E}[|f_{[\hat{m}]} - m|^2]\|_\infty \leq A(1 + DN^{\frac{-2\beta}{2\beta+1}})^2 + C^* N^{\frac{-2\beta}{2\beta+1}}, \tag{27}$$

where $A = \frac{144s^2\alpha^2(M^2+\sigma^2)}{(1-2s^2)C_1^2\eta^2\tau^2}$ and D depends on λ, α, β, C_1, η, τ, σ^2, and C^* depends on λ, α, β, C_1, η, σ^2. Moreover, the constant M can be estimated by $\tilde{M} = \max\{|y_0|, |y_1|, \cdots, |y_N|\}$.

The right-hand side of (27) tends to A when $N \to \infty$. In fact, if d is chosen so that $d \to 0$ and $Nd \to \infty$ as $N \to \infty$, then the right-hand side of (20) tends to A as $N \to \infty$. Moreover, $A \to 0$ as $s \to 0$.

Example 1. *The data set we used in this example is the Crude Oil WTI Futures daily highest price from 2021/7/19 to 2022/8/17. These data are opened and they can be obtained from the website https://www.investing.com/commodities/crude-oil-historical-data. There are 287 raw data and we chose 11 data as our sample subset $\mathcal{S}$. These data points are shown in Figure 1. We set $\mathcal{S} = \{(t_i, w_i) : i = 0, 1, \cdots, 10\}$, where $t_i = \frac{i}{10}$ and w_i are the Crude Oil WTI Futures daily highest prices in 2021/7/19, 8/26, 10/6, 11/16, 12/28, 2022/2/2, 3/9, 4/19, 5/30, 7/5, and 8/17.*

Let $\hat{m}$ be defined by (7), with each Y_i being replaced by w_i,

$$\hat{m}(t) = \frac{\sum_{i=0}^{10} k(\frac{t-0.1\times i}{d})w_i}{\sum_{j=0}^{10} k(\frac{t-0.1\times j}{d})}, \tag{28}$$

and choose k to be the Epanechnikov kernel $k(t) = 0.75(1-t^2)\chi_{\{|t|\leq 1\}}$. Let $N = 10$ and choose $R = 1$, $C_2 = 1$, $\eta = \frac{1}{\sqrt{3}}$, $C_1 = 0.5$ in (8). We estimate M by $\max\{w_0, w_1, \cdots, w_{10}\}$, and set $\alpha = 1$ and $\tau = 1$ in Theorem 3. Assume that $\beta = 0.5$ in this example. The values of σ^2 and λ are estimated by the sample variance and $\max\left\{\frac{|w_i - w_j|}{\sqrt{|t_i - t_j|}} : i, j = 0, 1, \cdots, 10, i \neq j\right\}$, respectively. By (17), we set $d = 0.092$.

We construct a FIF $f_{[\hat{m}]}$ by the method given in Section 2 with linear functions L_i and the linear polynomial p such that $L_i(0) = \frac{i-1}{10}$, $L_i(1) = \frac{i}{10}$, and $p(0) = \hat{m}(0)$, $p(1) = \hat{m}(1)$. The chosen values $s_1, \ldots, s_{10}$ are given in Table 1.

The graphs of raw data and $\hat{m}$ are shown in Figure 2. The graphs of $\hat{m}$ and $f_{[\hat{m}]}$ are shown in Figure 3. The graphs of raw data and $f_{[\hat{m}]}$ are shown in Figure 4.

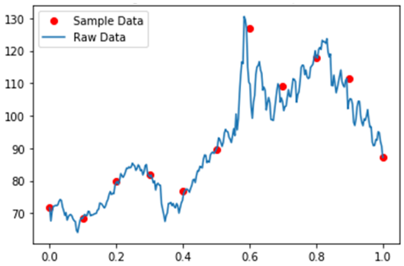

Figure 1. Raw data and sample data.

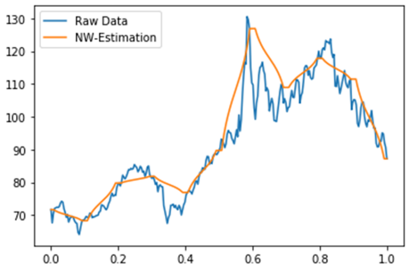

Figure 2. Raw data and $\hat{m}$.

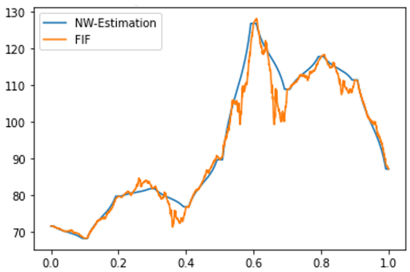

Figure 3. $\hat{m}$ and $f_{[\hat{m}]}$.

Figure 4. Raw data and $f_{[\hat{m}]}$.

Table 1. The values of s_k.

k	1	2	3	4	5	6	7	8	9	10
s_k	0.02	−0.03	0.08	−0.16	0.05	−0.26	−0.36	−0.06	−0.14	0.06

5. Conclusions

The purpose of this paper is to construct a fractal interpolation function (FIF) that has good approximation for a given data set. We consider the Nadaraya–Watson estimator $\hat{m}$ for some sample data chosen from a given data set, and then apply $\hat{m}$ to construct a FIF $f_{[\hat{m}]}$ to fit the given set of data points. The Nadaraya–Watson estimator is widely used in data-fitting problems, and its fractal perturbation is considered in our paper. The expectations of mean squared errors of such approximation are also estimated. By the figures given in Example 1, we may see the quality of curve fitting by a FIF, which is constructed from $\hat{m}$ with 11 sample points to fit the 287 raw data points. We see that the error of approximation can be decreased by choosing more sample data.

In this paper, we construct a FIF to fit a given data set with the help of the Nadaraya–Watson estimator. In fact, the Priestley–Chao estimator, the Gasser–Müller estimator, and other types of kernel regression estimators can also be used in our approach. Nonparametric regression has been studied for a long time. Several types of models with their theoretical results and applications are widely developed by many researchers. Fractal perturbations of these models are worth investigating in the field of fractal curve fitting.

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

Funding: This research was funded by Ministry of Science and Technology, R.O.C. grant number MOST 110-2115-M-214-002.

Data Availability Statement: The data set used in this paper can be obtained in the webpage https://www.investing.com/commodities/crude-oil-historical-data.

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

References

1. Györfi, L.; Kohler, M.; Krzyzak, A.; Walk, H. *A Distribution-Free Theory of Nonparametric Regression*; Springer: New York, NY, USA, 2002.
2. Härdle, W.; Müller, M.; Sperlich, S.; Werwatz, A. *Nonparametric and Semiparametric Models*; Springer: New York, NY, USA, 2004.
3. Hart, J.D. *Nonparametric Smoothing and Lack-of-Fit Tests*; Springer: New York, NY, USA, 1997.
4. Li, Q.; Racine, J.S. *Nonparametric Econometrics*; Princeton University Press: Mercer County, NJ, USA, 2007.
5. Tsybakov, A.B. *Introduction to Nonparametric Estimation*; Springer: New York, NY, USA, 2009.
6. Wasserman, L. *All of Nonparametric Statistics*; Springer: New York, NY, USA, 2006.
7. Chu, C.-K.; Marron, J.S. Choosing a kernel regression estimator. *Stat. Sci.* **1991**, *6*, 404–436. [CrossRef]
8. Jones, M.C.; Davies, S.J.; Park, B.U. Versions of kernel-type regression estimators. *J. Amer. Statist. Assoc.* **1994**, *89*, 825–832. [CrossRef]
9. Barnsley, M.F. Fractal functions and interpolation. *Constr. Approx.* **1986**, *2*, 303–329. [CrossRef]
10. Barnsley, M.F. *Fractals Everywhere*; Academic Press: Orlando, FL, USA, 1988.
11. Marvasti, M.A.; Strahle, W.C. Fractal geometry analysis of turbulent data. *Signal Process.* **1995**, *41*, 191–201. [CrossRef]
12. Mazel, D.S. Representation of discrete sequences with three-dimensional iterated function systems. *IEEE Trans. Signal Process.* **1994**, *42*, 3269–3271. [CrossRef]
13. Mazel, D.S.; Hayes, M.H. Using iterated function systems to model discrete sequences. *IEEE Trans. Signal Process.* **1992**, *40*, 1724–1734. [CrossRef]
14. Balasubramani, N. Shape preserving rational cubic fractal interpolation function. *J. Comput. Appl. Math.* **2017**, *319*, 277–295. [CrossRef]
15. Balasubramani, N.; Guru Prem Prasad, M.; Natesan, S. Shape preserving α-fractal rational cubic splines. *Calcolo* **2020**, *57*, 21. [CrossRef]
16. Barnsley, M.F.; Elton, J.; Hardin, D.; Massopust, P. Hidden variable fractal interpolation functions. *SIAM J. Math. Anal.* **1989**, *20*, 1218–1242. [CrossRef]
17. Barnsley, M.F.; Massopust, P.R. Bilinear fractal interpolation and box dimension. *J. Approx. Theory* **2015**, *192*, 362–378. [CrossRef]
18. Chand, A.K.B.; Kapoor, G.P. Generalized cubic spline fractal interpolation functions. *SIAM J. Numer. Anal.* **2006**, *44*, 655–676. [CrossRef]
19. Chand, A.K.B.; Navascués, M.A. Natural bicubic spline fractal interpolation. *Nonlinear Anal.* **2008**, *69*, 3679–3691. [CrossRef]
20. Chand, A.K.B.; Navascués, M.A. Generalized Hermite fractal interpolation. *Rev. Real Acad. Cienc. Zaragoza* **2009**, *64*, 107–120.
21. Chand, A.K.B.; Tyada, K.R. Constrained shape preserving rational cubic fractal interpolation functions. *Rocky Mt. J. Math.* **2018**, *48*, 75–105. [CrossRef]
22. Chand, A.K.B.; Vijender, N.; Viswanathan, P.; Tetenov, A.V. Affine zipper fractal interpolation functions. *BIT Numer. Math.* **2020**, *60*, 319–344. [CrossRef]

23. Chand, A.K.B.; Viswanathan, P. A constructive approach to cubic Hermite fractal interpolation function and its constrained aspects. *BIT Numer. Math.* **2013**, *53*, 841–865. [CrossRef]
24. Chandra, S.; Abbas, S.; Verma, S. Bernstein super fractal interpolation function for countable data systems. *Numer. Algorithms* **2022**. [CrossRef]
25. Dai, Z.; Wang, H.-Y. Construction of a class of weighted bivariate fractal interpolation functions. *Fractals* **2022**, *30*, 2250034. [CrossRef]
26. Katiyar, S.K.; Chand, A.K.B. Shape preserving rational quartic fractal functions. *Fractals* **2019**, *27*, 1950141. [CrossRef]
27. Katiyar, S.K.; Chand, A.K.B.; Kumar, G.S. A new class of rational cubic spline fractal interpolation function and its constrained aspects. *Appl. Math. Comput.* **2019**, *346*, 319–335. [CrossRef]
28. Luor, D.-C. Fractal interpolation functions with partial self similarity. *J. Math. Anal. Appl.* **2018**, *464*, 911–923. [CrossRef]
29. Massopust, P.R. *Fractal Functions, Fractal Surfaces, and Wavelets*; Academic Press: San Diego, CA, USA, 1994.
30. Massopust, P.R. *Interpolation and Approximation with Splines and Fractals*; Oxford University Press: New York, NY, USA, 2010.
31. Miculescu, R.; Mihail, A.; Pacurar, C.M. A fractal interpolation scheme for a possible sizeable set of data. *J. Fractal Geom.* **2022**. [CrossRef]
32. Navascués, M.A. Fractal approximation. *Complex Anal. Oper. Theory* **2010**, *4*, 953–974. [CrossRef]
33. Navascués, M.A. Fractal bases of L_p spaces. *Fractals* **2012**, *20*, 141–148. [CrossRef]
34. Navascués, M.A.; Chand, A.K.B. Fundamental sets of fractal functions. *Acta Appl. Math.* **2008**, *100*, 247–261. [CrossRef]
35. Navascués, M.A.; Pacurar, C.; Drakopoulos, V. Scale-free fractal interpolation. *Fractal Fract.* **2022**, *6*, 602. [CrossRef]
36. Prasad, S.A. Super coalescence hidden-variable fractal interpolation functions. *Fractals* **2021**, *29*, 2150051. [CrossRef]
37. Ri, S.; Drakopoulos, V. Generalized fractal interpolation curved lines and surfaces. *Nonlinear Stud.* **2021**, *28*, 427–488.
38. Tyada, K.R.; Chand, A.K.B.; Sajid, M. Shape preserving rational cubic trigonometric fractal interpolation functions. *Math. Comput. Simul.* **2021**, *190*, 866–891. [CrossRef]
39. Vijender, N. Fractal perturbation of shaped functions: convergence independent of scaling. *Mediterr. J. Math.* **2018**, *15*, 211. [CrossRef]
40. Viswanathan, P. A revisit to smoothness preserving fractal perturbation of a bivariate function: Self-Referential counterpart to bicubic splines. *Chaos Solitons Fractals* **2022**, *157*, 111885. [CrossRef]
41. Viswanathan, P.; Chand, A.K.B. Fractal rational functions and their approximation properties. *J. Approx. Theory* **2014**, *185*, 31–50. [CrossRef]
42. Viswanathan, P.; Chand, A.K.B. α-fractal rational splines for constrained interpolation. *Electron. Trans. Numer. Anal.* **2014**, *41*, 420–442.
43. Viswanathan, P.; Navascués, M.A.; Chand, A.K.B. Associate fractal functions in $\mathcal{L}^p$-spaces and in one-sided uniform approximation. *J. Math. Anal. Appl.* **2016**, *433*, 862–876. [CrossRef]
44. Wang, H.-Y.; Yu, J.-S. Fractal interpolation functions with variable parameters and their analytical properties. *J. Approx. Theory* **2013**, *175*, 1–18. [CrossRef]
45. Banerjee, S.; Gowrisankar, A. *Frontiers of Fractal Analysis Recent Advances and Challenges*; CRC Press: Boca Raton, FL, USA, 2022.
46. Kumar, M.; Upadhye, N.S.; Chand, A.K.B. Linear fractal interpolation function for data set with random noise. *Fractals* 2022, *accepted*. [CrossRef]
47. Luor, D.-C. Fractal interpolation functions for random data sets. *Chaos Solitons Fractals* **2018**, *114*, 256–263. [CrossRef]
48. Luor, D.-C. Statistical properties of linear fractal interpolation functions for random data sets. *Fractals* **2018**, *26*, 1850009. [CrossRef]
49. Luor, D.-C. Autocovariance and increments of deviation of fractal interpolation functions for random datasets. *Fractals* **2018**, *26*, 1850075. [CrossRef]
50. Luor, D.-C. On the distributions of fractal functions that interpolate data points with Gaussian noise. *Chaos Solitons Fractals* **2020**, *135*, 109743. [CrossRef]
51. Caldarola, F.; Maiolo, M. On the topological convergence of multi-rule sequences of sets and fractal patterns. *Soft Comput.* **2020**, *24*, 17737–17749. [CrossRef]
52. Wang, H.-Y.; Li, H.; Shen, J.-Y. A novel hybrid fractal interpolation-SVM model for forecasting stock price indexes. *Fractals* **2019**, *27*, 1950055. [CrossRef]
53. Tosatto, S.; Akrour, R.; Peters, J. An upper bound of the bias of Nadaraya-Watson kernel regression under Lipschitz assumptions. *Stats* **2021**, *4*, 1–17. [CrossRef]