axioms

MDPI

Article

Spatio-Functional Nadaraya–Watson Estimator of the Expectile Shortfall Regression

Mohammed B. Alamari [1], Fatimah A. Almulhim [2,*], Zoulikha Kaid [1] and Ali Laksaci [1]

[1] Department of Mathematics, College of Science, King Khalid University, Abha 62529, Saudi Arabia; malamari@kku.edu.sa (M.B.A.); zqayd@kku.edu.sa (Z.K.); alikfa@kku.edu.sa (A.L.)
[2] Department of Mathematical Sciences, College of Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
* Correspondence: faalmulhim@pnu.edu.sa

Abstract: The main aim of this paper is to consider a new risk metric that permits taking into account the spatial interactions of data. The considered risk metric explores the spatial tail-expectation of the data. Indeed, it is obtained by combining the ideas of expected shortfall regression with an expectile risk model. A spatio-functional Nadaraya–Watson estimator of the studied metric risk is constructed. The main asymptotic results of this work are the establishment of almost complete convergence under a mixed spatial structure. The claimed asymptotic result is obtained under standard assumptions covering the double functionality of the model as well as the data. The impact of the spatial interaction of the data in the proposed risk metric is evaluated using simulated data. A real experiment was conducted to measure the feasibility of the Spatio-Functional Expectile Shortfall Regression (SFESR) in practice.

Keywords: financial risk; complete consistency; expected shortfall; functional data; kernel method; expectile regression; quantile regression

MSC: 62G08; 62G10; 62G35; 62G07; 62G32; 62G30; 62H12

check for updates

Citation: Alamari, M.B.; Almulhim, F.A.; Kaid, Z.; Laksaci, A. Spatio-Functional Nadaraya–Watson Estimator of the Expectile Shortfall Regression. *Axioms* **2024**, *13*, 678. https://doi.org/10.3390/axioms13100678

Academic Editors: Christos Kountzakis, Konstantinos Gkillas and Christos Floros

Received: 12 August 2024
Revised: 17 September 2024
Accepted: 23 September 2024
Published: 30 September 2024

1. Introduction

Currently, the spatial correlation of data has a potential impact on financial risk management. Indeed, with the rapid development of internet technology, investors are increasingly interested in international financial assets, which requires taking into account the spatial dependence of international stock markets. Of course, unlike standard spatial data analysis, the spatial correlation in spatio-financial time series data is not necessarily measured by the geographic coordinates of the stock markets. This is the principal motivation for introducing a financial risk metric to cover the spatial component of risk management. Recall that spatial data cannot be treated as independent (see [1,2], among others). In practice, the challenging issue of spatial data analysis comes from the fact that points are in multi-dimensional space without linear order.

Statistical analysis of spatial data has become widely developed in the last decade. Concerning the nonparametric approach, the first results were obtained by the author of [3], who obtained the asymptotic normality for the density kernel estimator. The regression function was studied in [4,5], in which the authors employed an estimator from the Nadaraya–Watson weights techniques. We refer to [6] for the nonparametric kernel estimator for the variogram, considering Nadaraya–Watson weights. Ref. [7] investigated the local linear estimation for the regression function (see also [8] for the spatial auto-regression model) and proved the uniform convergence of the constructed estimator. Their convergence rate is optimal according to the L_∞-norm. In [9], we found an alternative local linear estimator of the spatial regression, which was obtained using the least absolute deviation. In this cited work, the authors have derived the asymptotic normality of their estimator.

We return to [10] for estimation using the nearest neighbor method. In functional statistics, the authors of [11] have constructed an estimator using the spatiotemporal process. They proved the almost complete convergence (a.co.) of their estimator when the input variable is a continuous time process. The spatial quantile regression was estimated by [12]. Their estimator was constructed by inverting the estimator of the cumulative distribution function. For a more bibliographic discussion of spatio-functional data analysis, we refer the reader to [13–16].

The second important component of this study is the shortfall function (ES). This is a risk management model and was created by [17]. The principal motivation of the expected shortfall function as a risk metric is its coherency property. The estimation of the ES model is performed using multiple algorithms such as parametric, nonparametric, or semi-parametric approaches. The recent advances and references on the parametric approaches can be found in [18–20]. While the nonparametric estimation was developed by [21], we also cite [22] for the functional Nadaraya–Watson estimator of the functional expected shortfall regression (FESR), in which the authors studied the asymptotic properties of FESR under the mixing assumption. The weak dependence case was treated by the authors of [23], who almost established complete consistency of the kernel estimator of the FESR using the quasi-associated structure. We point out that in previous studies, the expected loss in FSER is defined through the Value at Risk (VaR)-level, the so-called FSER-VaR. In this work, we introduce an alternative risk threshold defined by the expectile regression, the so-called FSER-expectile. The expectile regression is an alternative risk metric based on tail expectation, unlike the VaR function, which is based on tail frequency. For this reason, the use of the expectile instead of the VaR function is more informative because it is more sensitive to outliers. This feature increases its ability to fit the financial risk located in the extreme values. In recent years, the expectile model has gained popularity in risk analysis (see, for instance, [24–27] for more motivations for these models). Although previous studies focus on the unconditional models, in this paper, we focus on the regression case. This version of the expectile has been studied in multivariate statistics by many authors. The first results date back to [28]. In the last decade, multivariate expectile regression has been employed for many statistical issues, including additive models [29], neural network models [30], and machine learning models [27]. However, financial risk analysis seems to be the principal applied area of the expectile regression model. In this context, ref. [31] proposes an estimation of the value at risk (VaR) using an expectile model. Ref. [32] presents different approaches used to preserve the coherence properties of multivariate expectiles. The same authors in [33] established the asymptotic behavior of the multivariate expectiles for the Fréchet model. The treatment of the functional case was recently considered in [13], in which the authors considered expectile regression (ER) with a functional covariate. They constructed an estimator of the functional ER using the nonparametric kernel approach. An alternative approach was studied by the authors of [34] using the functional parametric ER. The authors employed a Hilbert structure using a reproducing kernel. More recent advances in functional expectile regression can be found in [15] and the references therein. We may return to [35–37] for more recent development in FTSA.

As discussed below, the main purpose of the present paper is to introduce a new risk metric based on the expectile shortfall regression. The developed risk metric has many advantages over the old shortfall model. These advantages are because the expectile is elicitable and coherent, unlike the VaR, and additionally, it is more sensitive to the magnitude of the tail, unlike the VaR function. Thus, the expectile shortfall with expectile (ESE) is more efficient than the standard shortfall. In this paper, we consider a more complex functional structure based on the spatial correlation. The spatial correlation is more general than the standard functional time series structure. It allows for controlling the spatial interaction of the data, which is more interactive in risk management. Furthermore, the principal outcomes of this work are the construction of a computational estimator and the establishment of its asymptotic properties using spatial dependence. The practical use of this risk metric is evaluated using simulated and real data. To the best of our knowledge, spatial expected shortfall regression has not yet been fully explored, and this is the first study in this direction.

This paper is organized as follows: We present our model as well as its spatial estimator in the next section. Section 2 is dedicated to introducing the spatio-functional time series framework. The almost complete convergence of the constructed estimator is shown in Section 3. Section 4 is devoted to examining the easy implementation of the estimator using simulated data. In Section 5, we apply our model to analyze the extreme values in environmental time series data. Some concluding remarks, as well as some future prospects, are discussed in Section 6. Finally, the proofs of the auxiliary results are given in Appendix A.

2. Model and Estimator

Consider $(A_{\mathbf{i}}, B_{\mathbf{i}})$, $\mathbf{i} \in \mathbb{Z}^N$, $N \geq 1$, a stationary spatial process defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and valued $\mathcal{F} \times \mathbb{R}$. $\mathcal{F}$ is a semi-metric space with d denoting the corresponding semi-metric. A point $\mathbf{i}$ will be referred to as a site and is defined by the components $(i_1, \ldots, i_N) \in \mathbb{Z}^N$. In this work, we focus on increasing domain asymptotic, where the underlining process, $(A_{\mathbf{i}}, B_{\mathbf{i}})$, is observed over a rectangular domain $I_{\mathbf{n}} = \left\{\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{Z}^N,\ 1 \leq i_k \leq n_k, k = 1, \ldots, N\right\}$, $\mathbf{n} = (n_1, \ldots, n_N) \in \mathbb{Z}^N$. Therefore, the index-vector $\mathbf{n} \to \infty$ means $\min\{n_k\} \to \infty$ and $|\frac{n_j}{n_k}| < C$ for all j, k such that $1 \leq j, k \leq N$ and for a given constant C such that $0 < C < \infty$. This kind of design is known as an asymptotically increasing domain, which allows the area of observations to become larger without large distances between the sites. Moreover, for $\mathbf{n} = (n_1, \ldots, n_N) \in \mathbb{Z}^N$, we set $\overline{\mathbf{n}} = \prod_{i=1}^N n_i$. The spectral structure of the functional random field $(A_{\mathbf{i}}, B_{\mathbf{i}})$, $\mathbf{i} \in \mathbb{Z}^N$, is controlled through the following mixing condition:

$$\begin{cases} \text{There exists a function } \psi(t) \downarrow 0 \text{ as } t \to \infty, \text{ such that} \\ \forall\, \mathcal{X},\ \mathcal{X}' \text{ subsets of } \mathbb{Z}^N \text{ has finite cardinals} \\ \alpha\Big(\mathcal{B}(\mathcal{X}), \mathcal{B}\Big(\mathcal{X}'\Big)\Big) = \sup_{B \in \mathcal{B}(\mathcal{X}), C \in \mathcal{B}(\mathcal{X}')} |\mathbb{P}(\mathbf{B} \cap \mathbf{C}) - \mathbb{P}(\mathbf{B})\mathbb{P}(\mathbf{C})| \\ \leq \phi\Big(\mathrm{Card}(\mathcal{X}), \mathrm{Card}(\mathcal{X}')\Big)\psi\Big(\mathrm{dist}\Big(\mathcal{X}, \mathcal{X}'\Big)\Big), \end{cases} \tag{1}$$

where $\mathcal{B}(\mathcal{X})$ (respectively, $\mathcal{B}(\mathcal{X}')$) means the Borel σ-field generated by $(A_{\mathbf{i}},\ \mathbf{i} \in \mathcal{X})$ (respectively, $\Big(A_{\mathbf{i}},\ \mathbf{i} \in \mathcal{X}'\Big)$), $\mathrm{Card}(\mathcal{X})$ (respectively, $\mathrm{Card}(\mathcal{X}')$) is the cardinality of $\mathcal{X}$ (respectively, $\mathcal{X}'$), $\mathrm{dist}\Big(\mathcal{X}, \mathcal{X}'\Big)$ is the Euclidean distance between $\mathcal{X}$ and $\mathcal{X}'$ and $\phi : \mathbb{Z}^2 \to \mathbb{R}^+$ is a symmetric positive function nondecreasing in each variable, such that $\forall n, m, \in \mathbb{Z}$

$$\phi(n, m) \leq C \min(n, m), \quad C > 0. \tag{2}$$

$$\sum_{i=1}^{\infty} i^{\delta} \psi(i) < \infty,\ \ \delta > 0. \tag{3}$$

Note that condition (2) can be replaced by

$$\phi(n, m) \leq C(n + m + 1)^{\widetilde{\beta}} \qquad \text{for some } \widetilde{\beta} > 1. \tag{4}$$

Both conditions (2) and (4) are used in Tran [3] and Carbon et al. [8], and are satisfied by many spatial models (see [38] for some examples). It should be noted that if $N = 1$, then $(A_{\mathbf{i}}, B_{\mathbf{i}})$ is called a strongly mixing process.

Throughout this paper, for a fixed point $\mathfrak{z} \in \mathcal{F}$, we denote by $\mathcal{N}_{\mathfrak{z}}$ for a given neighborhood of $\mathfrak{z}$. We assume that $(A_{\mathbf{i}}, B_{\mathbf{i}})$'s have the same distribution as (A, B). We put $CDF(\cdot|\mathfrak{z}')$, the conditional distribution of B given $A = \mathfrak{z}'$, and we assume the regular version of this conditional distribution exists for any $\mathfrak{z}' \in \mathcal{N}_{\mathfrak{z}}$. Additionally, we suppose that $CDF(\cdot|\mathfrak{z})$ has a continuous density $f(\cdot|\mathfrak{z})$ with respect to Lebesgue's measure over $\mathbb{R}$.

Recall that the standard FESR regression is defined

$$\text{for all } \mathfrak{z} \in \mathcal{F}, \text{ by} \quad RES_p(\mathfrak{z}) = \mathbb{E}\big[B|B > RVaR_p(\mathfrak{z}), A = \mathfrak{z}\big],$$

where $RVaR_p$ is the conditional quantile of order $1-p$. Clearly, it is defined through the tail quantile, which is frequency-tail. Alternatively, it would be more interesting to evaluate this metric using the expectation tail. To do that, we introduce the FESR-expectile defined

$$\text{for all } \mathfrak{z} \in \mathcal{F}, \text{ by} \quad REA_p(\mathfrak{z}) = \mathbb{E}\big[B|B > REXP_p(\mathfrak{z}), A = \mathfrak{z}\big],$$

where $REXP_p$ is

$$\text{the expectile regression} \quad REXP_p(\mathfrak{z}) = \arg\min_{t\in\mathbb{R}} \Big\{ \mathbb{E}\Big[p(B-t)^2 \mathbb{1}_{\{(B-t)>0\}} \mid A = \mathfrak{z}\Big] + \mathbb{E}\Big[(1-p)(B-t)^2 \mathbb{1}_{\{(B-t)\leq 0\}} \mid A = \mathfrak{z}\Big] \Big\},$$

where $\mathbb{1}_{\mathbf{C}}$ is the indicator function of the set $\mathbf{C}$. It should be noted that the replacement of $RVaR_p$ by $REXP_p$ is important in practice, as it permits remedying the lack of risk insensitivity of $RVaR_p$ to the extreme values.

Now, to estimate $REA_p(\mathfrak{z})$ using the kernel estimator, we consider $\mathbf{F}(\cdot)$, a measurable function, $r = r_{\mathbf{n}}$ a positive sequence of real numbers tending to zero as $\mathbf{n}$ tends to infinity, and we estimate the FESR-expectile by

$$\widehat{REA_p}(\mathfrak{z}) = \frac{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathbf{F}\Big[r^{-1} d(\mathfrak{z}, A_{\mathbf{i}})\Big] B_{\mathbf{i}} \mathbb{1}_{B_{\mathbf{i}} > \widehat{REXP}_p(\mathfrak{z})}}{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathbf{F}\Big(r^{-1} d(\mathfrak{z}, A_{\mathbf{i}})\Big)}, \tag{5}$$

where $\widehat{REXP}_p$ is the kernel estimator of $REXP_p$, defined as the solution of

$$\widehat{G}(\widehat{REXP}_p(t;\mathfrak{z})) = \frac{p}{1-p}$$

with

$$\widetilde{G}(t;\mathfrak{z}) = \frac{-\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathbf{F}_{ni}(\mathfrak{z})(B_{\mathbf{i}} - t)\mathbb{1}_{\{(B_{\mathbf{i}}-t)\leq 0\}}}{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathbf{F}_{ni}(\mathfrak{z})(B_{\mathbf{i}} - t)\mathbb{1}_{\{(B_{\mathbf{i}}-t)> 0\}}}, \quad \text{for } t \in \mathbb{R},$$

where

$$\mathbf{F}_{ni}(\mathfrak{z}) = \frac{\mathbf{F}\big[r^{-1} d(\mathfrak{z}, A_{\mathbf{i}})\big]}{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathbf{F}\Big[r^{-1} d(\mathfrak{z}, A_{\mathbf{i}})\Big]}.$$

We refer to [13] for more discussion on the construction of the estimator $\widehat{REXP}_p$. While the estimator $\widehat{REA_p}$ is constructed using similar ideas to those used for classical regression [39], it is clear that the choice of the parameter r is primordial in this smoothing approach. It is crucial for the estimation of $\widehat{REA_p}$ as well as for $\widehat{REXP}_p$. Motivated by the strong relationship between the expectile and the mean squared error (MSE), the MSE-based cross-validation criterion is an appropriate rule with which to address this issue. The latter is common in nonparametric functional data analysis:

$$r_{opt} = \arg\min_{r} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \Big(B_{\mathbf{i}} - \widehat{REXP}_{0.5}(A_{\mathbf{i}})\Big)^2. \tag{6}$$

The popularity of this approach comes from its easy implementation in real data analysis, using the fact that the conditional mean $\mathbb{E}[Y|X]$ is associated with $\widehat{REXP}_p$ with $p = 0.5$.

3. Main Asymptotic Result

Before stating the asymptotic properties of the estimator $\widehat{REA_p}$, we need to introduce some notations and assumptions. Firstly, we set $C_{\mathfrak{z}}$ or $C'_{\mathfrak{z}}$ as some strictly positive generic constants, and for all $t \in \mathbb{R}$, we define $ES(t, \mathfrak{z}) = \mathbb{E}[B\mathbb{1}_{B>t}|\, A = \mathfrak{z}]$. Now, to formulate our main results, we will use the hypotheses listed below:

(P1) $P(A \in B(\mathfrak{z}, r)) = \phi(\mathfrak{z}, r) > 0$ where $B(\mathfrak{z}, r) = \{x' \in \mathcal{F}\ : d(\mathfrak{z}', \mathfrak{z}) < r\}$.

(P2) $\exists \delta > 0, \forall (t_1, t_2) \in [REXP_p(\mathfrak{z}) - \delta,\ REXP_p(\mathfrak{z}) + \delta], \forall (\mathfrak{z}_1, \mathfrak{z}_2) \in N_x^2$,

$$|ES(t_1, \mathfrak{z}_1) - ES(t_2, \mathfrak{z}_2)| \le C_x \left(d^b(\mathfrak{z}_1, \mathfrak{z}_2) + |t_1 - t_2|^b \right), \quad b > 0.$$

(P3) The sequence $(A_{\mathbf{i}}, B_{\mathbf{i}})_{\mathbf{i} \in \mathcal{I}_{\mathbf{n}}}$ such that

$$\begin{cases} \forall \mathbf{i} \neq \mathbf{j}, \quad 0 < \sup_{\mathbf{i} \neq \mathbf{j}} \mathbb{P}\left[(A_{\mathbf{i}}, A_{\mathbf{j}}) \in B(x, r) \times B(x, r)\right] \le C_1 (\phi(\mathfrak{z}, r))^{(a+1)/a}, \\ \qquad\qquad \text{for some } 1 < a < \delta N^{-1}. \\ \forall t \in [\theta_x - \delta, \theta_x + \delta], \qquad \mathbb{E}\left[B_{\mathbf{i}} B_{\mathbf{j}} | A_{\mathbf{i}}, A_{\mathbf{j}}\right] \le C < \infty, \\ \mathbb{E}\left[|B|^2 | X\right] < C < \infty \quad \text{and} \quad \mathbb{E}\left[|B|^p\right] < C < \infty, \quad p > 1 \end{cases}$$

(P4) $\mathbf{F}$ is a function with support $(0, 1)$ such that
$0 < C\mathbb{I}_{(0,1)} < \mathbf{F}(t) < C'\mathbb{I}_{(0,1)} < \infty$.

(P5) There exists $\eta_0 > 0$, such that,

$$C\overline{\mathbf{n}}^{\frac{(b-1)N - b\delta}{b\delta} + \eta_0} \le \phi(\mathfrak{z}, r)$$

Comments on the hypotheses.
Hypothesis (P1) is checked for several continuous time processes (see, for instance, [40] for a general Gaussian process). The local dependency in the first part (P3) allows us to obtain the same convergence rate as in the i.i.d. case. These hypotheses could be weakened, but the convergence rate would be perturbed by the presence of covariance terms (see Liebscher [41]). (P3) is a mild regularity hypothesis imposed to evaluate the bias term. The assumptions (P4)–(P5) are technical conditions for simplifying the proofs.

Now, we obtain the convergence rate of the almost complete convergence (a.co.) of the estimator $\widehat{REA_p}(\mathfrak{z})$ to $REA_p(\mathfrak{z})$. This stochastic convergence is stronger than the convergence in probability and almost sure convergence.

Theorem 1. *Under the suppositions (P1)–(P5), we have*

$$\left|\widehat{REA_p}(\mathfrak{z}) - REA_p(\mathfrak{z})\right| = O\left(r^b\right) + O\left(\left(\frac{\ln \overline{\mathbf{n}}}{\overline{\mathbf{n}}\, \phi(\mathfrak{z}, r)} \right)^{1/2} \right) \ a.co. \ as \ \ \mathbf{n} \to \infty. \tag{7}$$

4. Simulated Data

In this section, we aim to evaluate the impact of the spatial dependency on the finite-sample performance of the spatio-functional expectile shortfall estimator. In order to highlight the main feature of our procedure, we compare its sensitivity to the volatility of the data in two situations (homoscedastic and heteroscedastic cases). For this purpose, we generate the data from the following regression relationship

$$\text{Model M1}: \quad Y_{\mathbf{i}} = \int_0^1 5cos((4 - A_{\mathbf{i}}(t))^2\pi)dt + \epsilon_{\mathbf{i}}, \quad \mathbf{i} = (i_1, i_2) \quad N = 2$$

$$\text{Model M2}: \quad Y_{\mathbf{i}} = \int_0^1 1.5exp(A_{\mathbf{i}}(t))dt + \left(\int (5\log((4 - A_{\mathbf{i}}(t))^2))dt\right)\epsilon_{\mathbf{i}}.$$

where $\epsilon_{\mathbf{i}}$ is a Gaussian random field that has an exponential covariogram function,

$$C(u) = \sigma^2 e^{\frac{-u}{\phi}} \quad u \in [0, \infty). \tag{8}$$

Now, in order to fit the financial risk management context, we draw the spatio-functional input variables using a spatial ARCH process. This consideration allows us to simulate the spatial interaction in the co-movement of stock markets. Indeed, let $R_{t,\mathbf{i}}$, the log-return of a financial asset at time t on the stock market $\mathbf{i}$, be generated by a spatial ARCH process

$$R_{t,\mathbf{i}} = \Sigma_{t,\mathbf{i}} Z_{t,\mathbf{i}},$$

where $Z_{t,\mathbf{i}}$ is a sequence of random variables that are independent in t and identically distributed with zero mean, unit variance, and constant covariance matrix C. The conditional variance $\Sigma_{t,\mathbf{i}}$ is defined by

$$\Sigma^2_{t,\mathbf{i}} = \alpha' + \rho \sum_{\mathbf{j}} w_{\mathbf{i},\mathbf{j}} P^2_{t-1,\mathbf{j}},$$

where $w_{\mathbf{i},\mathbf{i}}$ is a known Spatial Weight Matrix (SWM). In fact, this kind of spatio-functional process is obtained using the routine code `sim.spARCH` in the `R`-package `spGARCH`. A sample of the functional co-variate is plotted in Figure 1.

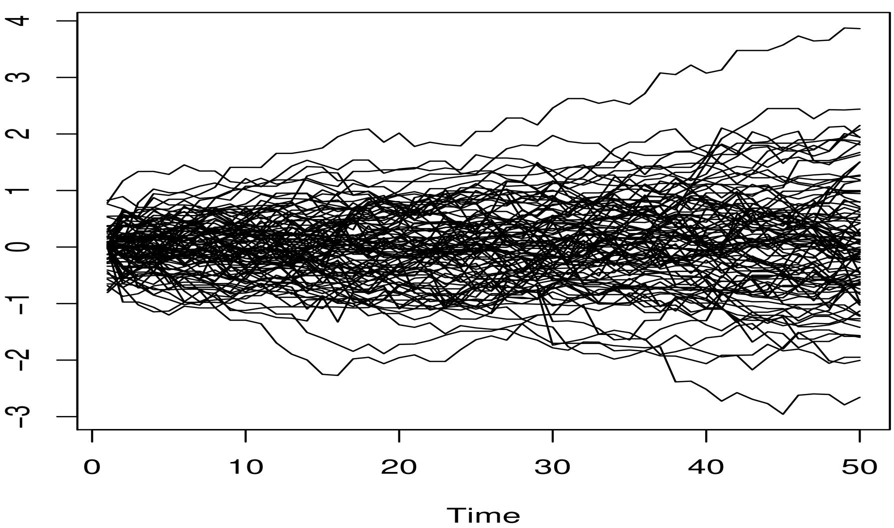

Figure 1. The ARCH process for $\alpha' = 0.05$ and $\rho = 0.8$.

Recall that the principal feature of the FESR-expectile is its high sensitivity to the outliers. To measure the impact of this characteristic, we use the routine code *ODM* in the R-Package *OutlierDM* to detect the number of outliers in each model. It appears that the first model contains 4% versus 28% for the second one. On the other hand, the spatial-heterogeneity of the data constitutes a second principal issue of our study. The latter is controlled through the parameters σ, ϕ and the spatial weight matrix $w_{\mathbf{i},\mathbf{i}}$. So, we calculate

$$MSE(p) = \bar{\mathbf{n}}^{-1} \sum_{\mathbf{i} \in \mathbf{I_n}} \left(B_{\mathbf{i}} - \widehat{REA_p}(A_{\mathbf{i}})\right)^2 \mathbb{1}_{B_{\mathbf{i}} > \widehat{REA_p}(A_{\mathbf{i}})}$$

for various values of the mentioned parameters.

Now, for this empirical study, we choose the smoothing parameter r via the local mean square cross-validation method as in (6). In the sense that the optimization of the mean square rule is performed over a discrete set defined by the k^{th}-distance from the location point. The integer number k is obtained from $\{5, 10, 15, 20, 25, 30, \ldots 50\}$. For the kernel $\mathbf{F}$, we use the β-kernel. Finally, the metric is chosen according to the nature of the functional variable and its smoothing property. It appears that the principal component (pca) metric is more suitable for this type of discontinuous functional regressor.

The simulation results are given in Table 1.

Table 1. Comparison results.

Model	n1	n2	σ	ϕ	SWM	MSE (0.01)	MSE (0.05)	MSE (0.5)	MSE (0.90)
M1	20	50	0.09	0.03	Queen	0.023	0.018	0.014	0.026
	50	30	0.09	0.03	Bishop	0.034	0.027	0.018	0.032
	20	30	0.79	0.93	Bishop	0.042	0.032	0.026	0.045
	20	50	0.09	0.03	Rook	0.042	0.020	0.018	0.037
	50	30	0.79	0.03	Rook	0.021	0.016	0.022	0.031
M2	20	50	0.09	0.03	Queen	0.045	0.036	0.028	0.044
	50	30	0.09	0.03	Bishop	0.071	0.053	0.026	0.059
	20	30	0.75	0.93	Bishop	0.096	0.052	0.048	0.105
	20	50	0.09	0.03	Rook	0.086	0.054	0.032	0.049
	50	30	0.79	0.03	Rook	0.039	0.025	0.047	0.055

We observe that the behavior of the estimator $\widehat{REA_p}$ is strongly affected by the different parameters of this study, such as the rate of the outliers and the spatial dependency degree. The high variability of the error between these different situations highlights the importance of the FESR- expectile as a risk-metric. In particular, the MSE varies between 0.018 and 0.045 with respect to the spatial level, while the horizontal variability, which describes the sensitivity to the outliers rate, ranges between 0.018 and 0.095.These results incorporate the theoretical study, where the convergence rate is strongly affected by the local dependency of the spatio-functional data. In the sense that the computational part proves that the performance of the estimator is strongly impacted by the degree of spatial correlation of the data. Such a conclusion highlights the importance of the expectile-based-shortfall. The latter is very sensitive to the variability or deviation of the data, allowing more reliability in risk detection. This feature makes the expectile-based-shortfall more appropriate as a risk metric than the standard expected shortfall. We point out that the standard expected shortfall is based on the quantile, which is a robust model with low sensitivity to the variability in the risk analysis, because the risk is often located in the extremes. Such a characteristic is not beneficial in risk analysis. Finally, we can say that the estimator $\widehat{REA_p}$ is very easy to implement and has good performance according to the nature of the treated data.

5. Real Data Application

After demonstrating the straightforward implementation of the estimator in the last section, we now focus on the applicability of our model to real spatial time series data. More specifically, we compare the performance of the new FESR-expectile $\widehat{REA_p}$ to the classical one

$$\widetilde{RES}_p(s) = \frac{\sum_{\mathbf{i}\in\mathbf{I_n}} \mathbf{F}(r^{-1}d(\mathfrak{z}, A_{\mathbf{i}})\Big(aG(a^{-1}(\widehat{RVaR}_p(\mathfrak{z}) - B_{\mathbf{i}})) + B_{\mathbf{i}}\Big(1 - H(a^{-1}\Big(\widehat{RVaR}_p(\mathfrak{z}) - B_{\mathbf{i}})\Big)\Big)\Big)}{p\sum_{\mathbf{i}\in\mathbf{I_n}} \mathbf{F}(r^{-1}d(\mathfrak{z}, A_{\mathbf{i}}))},$$

where $G(s) = \int_s^\infty u\mathbf{F}(u)$ and $H(s) = \int_{-infty}^s \mathbf{F}(u)du$. In the previous section, we evaluated the impact of spatial correlation using the ARCH model, which is well-solicited as an appropriate method for fitting the financial time series data. Alternatively, in this part, we employ the FESR-expectile model for another area, specifically in the environmental domain. This application emphasizes the importance and versatility of the FESR model.

The environmental domain is a particularly relevant area for risk management, as air quality significantly affects the quality of life. Moreover, the extreme values models have usually been employed to model the risk in this area. Here, we aim to compare the efficiency of the FESR- expectile $\widehat{REA_p}$ with the FESR-VaR $\widetilde{RES_p}$ in terms of risk prevention in air quality domain. For this goal, we analyze the air quality data used by [42], which concerns the ozone concentration in Beijing. These data are available on the website https://dataverse.harvard.edu/dataverse/beijing-air (accessed on 8 August 2024). Furthermore, there are many indices of air quality, such as Ozone (O_3), Particulate Matter (PM2.5 and PM10), Nitrogen Dioxide (NO_2), Carbon, and Sulfur Dioxide (SO_2). However, in this section, we concentrate on the ozone quantity (O_3) and sulfur dioxide (SO_2). Recall that the (SO_2) and the ultraviolet rays have a significant impact on the stratospheric ozone. Specifically, we collect the data from 120 monitoring stations in Beijing and we define $A_{\mathbf{i}}$ as the daily curve of SO_2 at the station $\mathbf{i}$ (on 30 December 2016). The response variable $B_{\mathbf{i}}$ represents the total ozone measured the day before at the same station $\mathbf{i}$. The daily curves for the sulphur dioxide are shown in Figure 2.

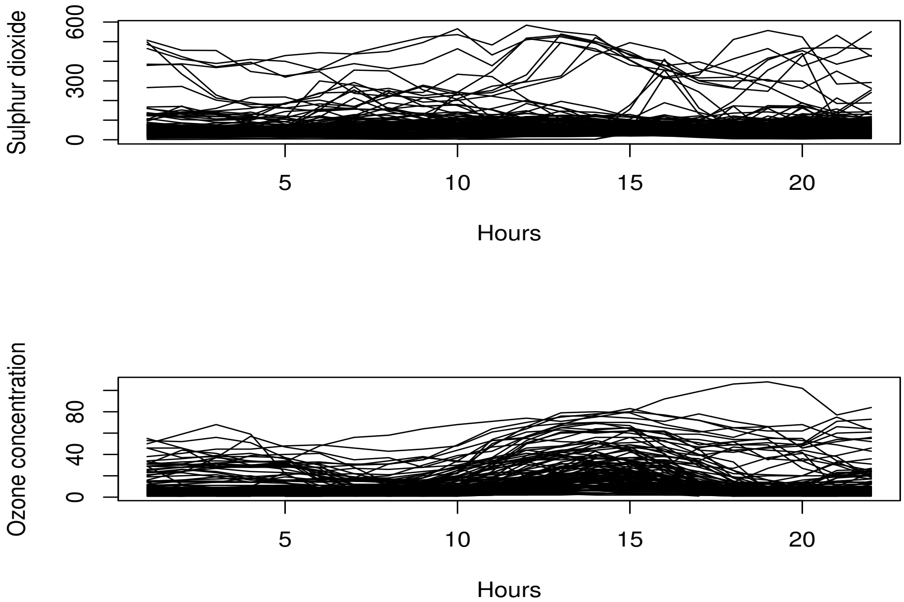

Figure 2. The SO_2 and O_3 daily curves.

Now, in order to explore the spatial correlation of the data, we follow the same strategy considered by [43]. This strategy permits us to estimate the spatial trend using the classical regression as follows. Indeed, we define

$$\widetilde{A}_{\mathbf{i}} = r_1(\mathbf{i}) + A_{\mathbf{i}} \quad \text{and} \quad \widetilde{B}_{\mathbf{i}} = r_2(\mathbf{i}) + B_{\mathbf{i}}.$$

Therefore, before computing the estimators $\widehat{REA_p}$ and $\widetilde{RES_p}$, we start by estimating the statistics $(\widehat{A}_{\mathbf{i}}, \widehat{B}_{\mathbf{i}})_{\mathbf{i}}$. The latter is estimated by

$$\widehat{A}_{\mathbf{i}} = \widetilde{A}_{\mathbf{i}} - \widehat{r}_1(\mathbf{i}) \quad \text{and} \quad \widehat{B}_{\mathbf{i}} = \widetilde{B}_{\mathbf{i}} - \widehat{r}_2(\mathbf{i}),$$

where $\widehat{r}_1(.)$ and $\widehat{r}_2(.)$ are the kernel estimators of the functions r_1 and r_2 which are

$$\widehat{m}_1(\mathbf{i}_0) = \frac{\sum_{\mathbf{i}\in\mathbf{I_n}} \mathbf{F}_1(r^{-1}\|\mathbf{i}_0 - \mathbf{i}\|)A_{\mathbf{i}}}{\sum_{\mathbf{i}\in\mathbf{I_n}} \mathbf{F}_1(r^{-1}\|\mathbf{i}_0 - \mathbf{i}\|)} \quad \left(\text{resp.} \quad \widehat{m}_2(\mathbf{j}_0) = \frac{\sum_{\mathbf{j}\in\mathbf{I_n}} \mathbf{F}_2(r^{-1}\|\mathbf{j}_0 - \mathbf{j}\|)B_{\mathbf{j}}}{\sum_{\mathbf{i}\in\mathbf{I_n}} \mathbf{F}_2(r^{-1}\|\mathbf{j}_0 - \mathbf{j}\|)}\right),$$

where $\mathbf{F}_1, \mathbf{F}_2$ are kernel functions. Such estimators are obtained using the routine code `npreg` in the R-package `np` with $\mathbf{F}_1 = \mathbf{F}_2$ being the quadratic kernel. This step is fundamental for spatio-functional data analysis and is referred to as the detrending step. To highlight the potential impact of spatial correlation, we compare our expected shortfall to the standard one in both cases: with or without detrending. Specifically, the estimation with detrending is calculated by $(\widehat{A}_{\mathbf{i}}, \widehat{B}_{\mathbf{i}})_{\mathbf{i}}$, while in the other case (without detrending), we use the initial observation $(A_{\mathbf{i}}, B_{\mathbf{i}})_{\mathbf{i}}$ to compute the estimators.

Furthermore, to calculate both estimators, we follow the same procedures used in the simulation section. In other words, we use the $(0,1)$ quadratic kernel and the pca-metric, along with local cross-validation for the bandwidth parameter. The efficiency of both estimators is evaluated by computing

$$MSE(p) = \overline{\mathbf{n}}^{-1} \sum_{\mathbf{i} \in \mathbf{I_n}} \left(B_{\mathbf{i}} - \widehat{\Theta_p}(A_{\mathbf{i}}) \right)^2 \mathbb{1}_{B_{\mathbf{i}} > \widehat{REX_p}(A_{\mathbf{i}})},$$

where $\widehat{\Theta_p}$ represents $\widehat{REA_p}$ or $\widetilde{RES_p}$. The values of $MSE()$ are evaluated as a function of p. In Figures 3 and 4, we show the values of MSE of both estimators $\widehat{REA_p}(black\ line)$ and $\widetilde{RES_p}(red\ line)$ in both cases (with detrending and without detrending step—see Figures 3 and 4).

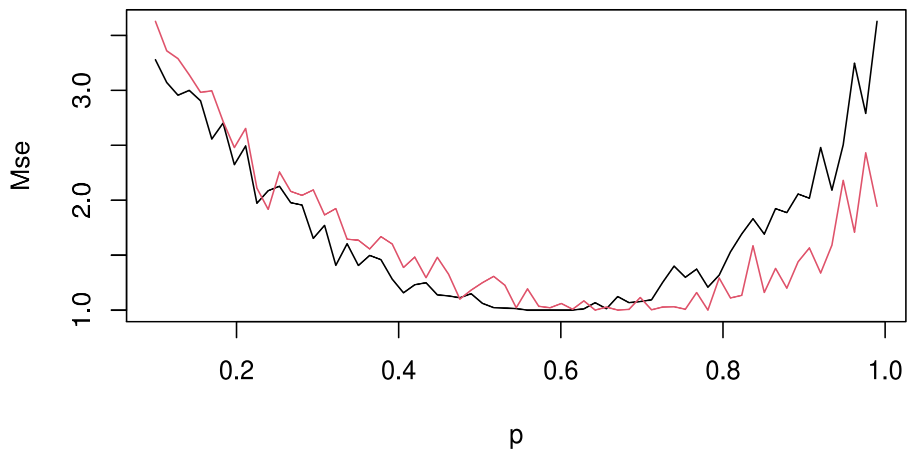

Figure 3. Comparison of the MSE values between FESR-expectile and FESR-VaR without detrending cases. The black line represents $\widehat{REA_p}$, and the red line represents $\widetilde{RES_p}$.

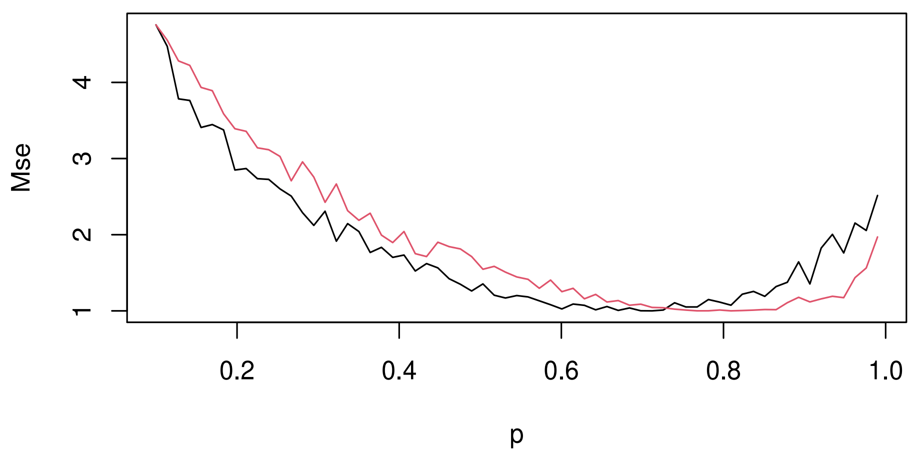

Figure 4. Comparison of the MSE values between FESR-expectile and FESR-VaR with detrending cases. The black line represents $\widehat{REA_p}$, and the red line represents $\widetilde{RES_p}$.

The graphs show the superiority of the FESR-expectile regression over the FESR-quantile model. This statement can be confirmed by the position of the black line, which is under the red line in most cases. These results show that the FESR-expectile detects the excessive level of ozone concentration more effectively, even in cases of high variability.

This feature is not surprising. The slow variability of the VaR level is due to the robustness of the quantile regression, which reduces its sensitivity to extreme values. Additionally, this advantage seems to be more significant in the detrending step compared to the non-detrending case. This statement can be confirmed using the cover test developed by Bayer and Dimitriadis [44]. This test allows us to examine the goodness-of-fit of our approach. The proposed test is an alternative approach to the procedure introduced by [45] for forecasting. Since the risk prediction differs significantly from standard prediction, we have opted to examine the feasibility of our risk-metric using the Bayer–Dimitriadis test. Specifically, we compare both functional approaches $\widetilde{RES_p}$ and $\widehat{REA_p}$ using the routine code esr-backtest from the R-package esrback. We have employed this code with $\alpha = 0.05$. Unsurprisingly, the obtained results confirm that both models are significantly good for this risk management issue. Typically, the cover-test gives a p-value of $\widehat{REA_p}$ equal to 0.001, compared to 0.004 for the model $\widetilde{RES_p}$.

6. Conclusions and Prospects

In this contribution, we have considered the nonparametric estimation of the FESR-regression-expectile under the spatial structure. We have constructed the functional version of the kernel estimator of this model as a risk-metric. This study covers a more general case of the functional random field. In the theoretical part, we have established the Borell–Contelli convergence under strong spatial mixing assumptions. Such theoretical development provides indispensable mathematical support for the use of the newly developed risk-metric. Additionally, the obtained asymptotic result was derived under general conditions and with the precision of the pointwise convergence rate. The computational part shows the applicability of the estimator and its very easy implementation in practice. Additionally, we applied the new model to an environmental spatio-functional random process. The result confirms the superiority of the FESR-expectile over FESR-VaR. On the other hand, the importance of this contribution can be viewed through several open future directions. For instance, we will address more dependent cases, such as the quasi-associated spatio-functional time series. This situation allows us to control the co-movement of different stock exchanges using weak dependence. The second issue is determining the uniform UNN convergence of the estimator, which will help in resolving the smoothing parameter selection. Furthermore, we can also estimate the model using either the additive or the linear case.

Author Contributions: The authors contributed approximately equally to this work. Formal analysis, A.L.; Validation, M.B.A.; Writing—review & editing, Z.K. and F.A.A. All authors have read and agreed to the final version of the manuscript.

Funding: This research is funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R515), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia; and the Deanship of Scientific Research and Graduate Studies at King Khalid University through the Research Groups Program under grant number R.G.P./128/45.

Data Availability Statement: The data used in this study are available through the link https://dataverse.harvard.edu/dataverse/beijing-air (accessed on 8 August 2024).

Acknowledgments: The authors thank and extend their appreciation to the funders of this work: This work was supported by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R515), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, and the Deanship of Scientific Research and Graduate Studies at King Khalid University through the Research Groups Program under grant number R.G.P./128/45.

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

Appendix A

This appendix is dedicated to proving the mathematical results of the paper.

Proof of the Theorem 1. We start by writing, for all $t \in \mathbb{R}$,

$$\widehat{ES}(t,\mathfrak{z}) = \frac{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathbf{F}\left[r^{-1}d(\mathfrak{z},A_{\mathbf{i}})\right] B_{\mathbf{i}}\mathbb{1}_{B_{\mathbf{i}}>t}}{\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \mathbf{F}\left[r^{-1}d(\mathfrak{z},A_{\mathbf{i}})\right]}.$$

Thus,

$$\widehat{ES}(\widehat{REXP}_p(\mathfrak{z}),\mathfrak{z}) = \widehat{REA_p}(\mathfrak{z}), \quad \text{and} \quad ES(REXP_p(\mathfrak{z}),\mathfrak{z}) = REA_p(\mathfrak{z}).$$

So,

$$\widehat{REA_p}(\mathfrak{z}) - REA_p(\mathfrak{z}) = \widehat{ES}(\widehat{REXP}_p(\mathfrak{z}),\mathfrak{z}) - ES(\widehat{REXP}_p(\mathfrak{z}),\mathfrak{z})$$
$$+ES(\widehat{REXP}_p(\mathfrak{z}),\mathfrak{z}) - ES(REXP_p(\mathfrak{z}),\mathfrak{z}).$$

Then,

$$|\widehat{REA_p}(\mathfrak{z}) - REA_p(\mathfrak{z})| \leq \sup_{t\in[REXP_p(\mathfrak{z})-\delta,\, REXP_p(\mathfrak{z})+\delta]} |\widehat{ES}(t,\mathfrak{z}) - ES(t,\mathfrak{z})|$$
$$+C|\widehat{REXP}_p(\mathfrak{z}) - REXP_p(\mathfrak{z})|.$$

So, the convergence rate in Theorem 1 is consequence of

$$\sup_{t\in[REXP_p(\mathfrak{z})-\delta,\, REXP_p(\mathfrak{z})+\delta]} |\widehat{ES}(t,\mathfrak{z}) - ES(t,\mathfrak{z})| = O\left(r^b\right) + O\left(\left(\frac{\ln \overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)^{1/2}\right) \qquad a.co. \tag{A1}$$

and

$$|\widehat{REXP}_p(\mathfrak{z}) - REXP_p(\mathfrak{z})| = O\left(r^b\right) + O\left(\left(\frac{\ln \overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)^{1/2}\right) \qquad a.co. \tag{A2}$$

As (A2) is proved in [13], it suffices to establish (A1). For this, we have $\mathbb{E}\left[\widehat{ES}_D(\mathfrak{z})\right] = 1$ and we write, for $t \in \mathbb{R}$

$$\widehat{ES}(t,\mathfrak{z}) - \widehat{ES}(t,\mathfrak{z}) = \frac{1}{\widehat{ES}_D(\mathfrak{z})}\left[\left(\widehat{ES}_N(t,\mathfrak{z}) - \mathbb{E}\left[\widehat{ES}_N(t,\mathfrak{z})\right]\right)\right.$$
$$\left.-\left(\widehat{ES}(t,\mathfrak{z})) - \mathbb{E}\left[\widehat{ES}_N(t,\mathfrak{z})\right]\right)\right] - \frac{\widehat{ES}_N(t,\mathfrak{z})}{\widehat{ES}_D(\mathfrak{z})}\left[\widehat{ES}_D(\mathfrak{z}) - \mathbb{E}\left[\widehat{ES}_D(\mathfrak{z})\right]\right].$$

Finally, the proof is a consequence of Lemmas A1–A3. □

Lemma A1. *Under the suppositions (P1) and (P3)–(P5), we have*

$$\widehat{ES}_D(\mathfrak{z}) - \mathbb{E}\left[\widehat{ES}_D(\mathfrak{z})\right] = O\left(\frac{\ln \overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)^{1/2} \qquad a.co.$$

Additionally,

$$\sum_n \mathbb{P}\left(\widehat{ES}_D(\mathfrak{z}) < \frac{1}{2}\right) < \infty.$$

Proof of Lemma A1. To prove this lemma, we use the classical spatial block decomposition (see [3]). Set

$$\widehat{ES}_D(\mathfrak{z}) = \frac{1}{\overline{\mathbf{n}}}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \frac{\mathbf{F}[r^{-1}d(\mathfrak{z},A_{\mathbf{i}})]}{\mathbb{E}[\mathbf{F}[r^{-1}d(\mathfrak{z},A_{\mathbf{i}})]]}.$$

We put $F_{\mathbf{i}} = \mathbf{F}\left[r^{-1}d(\mathfrak{z}, A_{\mathbf{i}})\right]$ and $D_{\mathbf{i}} = F_{\mathbf{i}} - \mathbb{E}[F_{\mathbf{i}}]$

$$\widehat{ES}_D(\mathfrak{z}) - 1 = \frac{1}{\overline{\mathbf{n}}\mathbb{E}[F_1]} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} D_{\mathbf{i}}.$$

So, we consider a sequence $p_{\mathbf{n}}$ and decompose the sum into 2^N partial sums of random variables as follows:

$$\mathcal{Y}(1, \mathbf{n}, x, \mathbf{j}) = \sum_{i_k=2j_kp_{\mathbf{n}}+1\ k=1,\ldots,N}^{2j_kp_{\mathbf{n}}+p_{\mathbf{n}}} D_{\mathbf{i}},$$

$$\mathcal{Y}(2, \mathbf{n}, x, \mathbf{j}) = \sum_{i_k=2j_kp_{\mathbf{n}}+1\ k=1,\ldots,N-1}^{2j_kp_{\mathbf{n}}+p_{\mathbf{n}}} \quad \sum_{i_N=2j_Np_{\mathbf{n}}+p_{\mathbf{n}}+1}^{(j_N+1)p_{\mathbf{n}}} D_{\mathbf{i}},$$

$$\mathcal{Y}(3, \mathbf{n}, x, \mathbf{j}) = \sum_{i_k=2j_kp_{\mathbf{n}}+1\ k=1,\ldots,N-2}^{2j_kp_{\mathbf{n}}+p_{\mathbf{n}}} \quad \sum_{i_{N-1}=2j_{N-1}p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(j_{N-1}+1)p_{\mathbf{n}}} \quad \sum_{i_N=2j_Np_{\mathbf{n}}+1}^{2j_Np_{\mathbf{n}}+p_{\mathbf{n}}} D_{\mathbf{i}},$$

$$\mathcal{Y}(4, \mathbf{n}, x, \mathbf{j}) = \sum_{i_k=2j_kp_{\mathbf{n}}+1\ k=1,\ldots,N-2}^{2j_kp_{\mathbf{n}}} \quad \sum_{i_{N-1}=2j_{N-1}p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(j_{N-1}+1)p_{\mathbf{n}}} \quad \sum_{i_N=2j_Np_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(j_N+1)p_{\mathbf{n}}} D_{\mathbf{i}},$$

and so on. Finally

$$\mathcal{Y}(2^{N-1}, \mathbf{n}, x, \mathbf{j}) = \sum_{i_k=2j_kp_{\mathbf{n}}+p_{\mathbf{n}}+1\ k=1,\ldots,N-1}^{2(j_k+1)p_{\mathbf{n}}} \quad \sum_{i_N=2j_Np_{\mathbf{n}}+1}^{2j_Np_{\mathbf{n}}+p_{\mathbf{n}}} D_{\mathbf{i}},$$

$$\mathcal{Y}(2^N, \mathbf{n}, x, \mathbf{j}) = \sum_{i_k=2j_kp_{\mathbf{n}}+p_{\mathbf{n}}+1\ k=1,\ldots,N}^{2(j_k+1)p_{\mathbf{n}}} D_{\mathbf{i}}.$$

Setting
$\mathcal{J} = \{0, \ldots, r_1 - 1\} \times \cdots \times \{0, \ldots, r_N - 1\}$, where $r_i = 2n_ip_{\mathbf{n}}^{-1}$, $i = 1, \ldots, N$ and we denote by

$$T(\mathbf{n}, x, i) = \sum_{\mathbf{j}\in\mathcal{J}} \mathcal{Y}(i, \mathbf{n}, x, \mathbf{j}).$$

Now, we write,

$$|\widehat{ES}_D(\mathfrak{z}) - \mathbb{E}[\widehat{ES}_D(\mathfrak{z})]| = \frac{1}{\overline{\mathbf{n}}\mathbb{E}[F_1]} \sum_{i=1}^{2^N} T(\mathbf{n}, x, i).$$

As regards this last inequality, we have $\forall\, \eta > 0$

$$\mathbb{P}\Big(|\widehat{ES}_D(\mathfrak{z}) - \mathbb{E}[\widehat{ES}_D(\mathfrak{z})]| \geq \eta\Big) \leq 2^N \max_{i=1,\ldots} \mathbb{P}(T(\mathbf{n}, x, i) \geq \eta\overline{\mathbf{n}}\mathbb{E}[F_1]).$$

Finally,

$$\mathbb{P}(T(\mathbf{n}, x, i) \geq \eta\overline{\mathbf{n}}\mathbb{E}[F_1]), \qquad \text{for all } i = 1, \ldots, 2^N.$$

We enumerate the $M = \prod_{k=1}^{N} r_k = 2^{-N}\overline{\mathbf{n}} p_{\mathbf{n}}^{-N} \leq \overline{\mathbf{n}} p_{\mathbf{n}}^{-N}$ random variables $\mathcal{Y}(1, \mathbf{n}, x, \mathbf{j})$; $\mathbf{j} \in \mathcal{J}$ in the arbitrary way $X_1, \ldots X_M$. Thus, for each X_j there exists a certain $\mathbf{j}_j$ in $\mathcal{J}$ such that

$$X_j = \sum_{\mathbf{i} \in I(1, \mathbf{n}, x, \mathbf{j}_j)} D_{\mathbf{i}},$$

where $I(1, \mathbf{n}, x, \mathbf{j}_j) = \left\{\mathbf{i} : 2j_{k_j} p_{\mathbf{n}} + 1 \leq i_k \leq 2j_{k_j} p_{\mathbf{n}} + p_{\mathbf{n}} \quad ; k = 1, \ldots N\right\}$. Observe that these sets contain $p_{\mathbf{n}}^N$ sites and are far apart by the distance of $p_{\mathbf{n}}^N$.

Now, we apply Lemma [8]. It permits the approximation of X_1,X_2, …, X_M by some independent random variables $X_1^*, \ldots X_M^*$, which have the same low as $X_{j=1,\ldots M}$, and such that

$$\sum_{j=1}^{r} \mathbb{E}|X_{\mathbf{j}} - X_{\mathbf{j}}^*| \leq 2CMp_{\mathbf{n}}^N \phi((M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N)\psi(p_{\mathbf{n}}).$$

So, we have to evaluate $\mathbb{P}(T(\mathbf{n}, x, 1) \geq \eta)$. For that, we employ Bernstein and Markov inequalities that

$$\mathbb{P}(T(\mathbf{n}, x, i) \geq \eta \overline{\mathbf{n}} \mathbb{E}[F_{\mathbf{1}}]) \leq \mathcal{B}_1 + \mathcal{B}_2$$

where

$$\mathcal{B}_1 = \mathbb{P}\left(\left|\sum_{j=1}^{M} X_{\mathbf{j}}^*\right| \geq \frac{M\eta\overline{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}]}{2M}\right) \leq 2\exp\left(-\frac{(\eta\overline{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}])^2}{MVar\left[X_1^*\right] + Cp_{\mathbf{n}}^N \eta\overline{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}]}\right)$$

and

$$\begin{aligned}
\mathcal{B}_2 &= \mathbb{P}\left(\sum_{j=1}^{M} |X_{\mathbf{j}} - X_{\mathbf{j}}^*| \geq \frac{\eta\overline{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}]}{2}\right) \\
&\leq \frac{1}{\eta\overline{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}]} \sum_{j=1}^{M} \mathbb{E}|X_{\mathbf{j}} - X_{\mathbf{j}}^*| \\
&\leq 2Mp_{\mathbf{n}}^N (\eta\overline{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}])^{-1} \phi((M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N)\psi(p_{\mathbf{n}}).
\end{aligned}$$

Since $\overline{\mathbf{n}} = 2^N M p_{\mathbf{n}}^N$ and $\phi((M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N) \leq p_{\mathbf{n}}^N$, we get for $\eta = \eta_0 \sqrt{\frac{\ln \overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z}, r)}}$

$$\mathcal{B}_2 \leq \overline{\mathbf{n}} p_{\mathbf{n}}^N (\ln \overline{\mathbf{n}})^{-1/2} (\overline{\mathbf{n}}\phi(\mathfrak{z}, r))^{-1/2} \varphi(p_{\mathbf{n}}).$$

As $p_{\mathbf{n}} = C\left(\frac{\overline{\mathbf{n}}\phi(\mathfrak{z}, r)}{\ln \overline{\mathbf{n}}}\right)^{1/2N}$, we write

$$\mathcal{B}_2 \leq \overline{\mathbf{n}}\, \psi(p_{\mathbf{n}}). \tag{A3}$$

Consequently, from (P5), we have

$$\Sigma_{\mathbf{n}} \overline{\mathbf{n}}\, \psi(p_{\mathbf{n}}) < \infty.$$

Let us focus now on $\mathcal{B}_1$. Indeed,

$$Var[X_1^*] = Var\left[\Sigma_{\mathbf{i} \in I(1, \mathbf{n}, x, \mathbf{1})} D_{\mathbf{i}}\right] = \Sigma_{\mathbf{i}, \mathbf{j} \in I(1, \mathbf{n}, x, \mathbf{1})} \left|Cov(D_{\mathbf{i}}, D_{\mathbf{j}})\right|.$$

Let $Q_{\mathbf{n}} = \Sigma_{\mathbf{i} \in I(1, \mathbf{n}, x, \mathbf{1})} Var[D_{\mathbf{i}}]$ and $\mathcal{R}_{\mathbf{n}} = \Sigma_{\mathbf{i} \neq \mathbf{j} \in I(1, \mathbf{n}, x, \mathbf{1})} \left|Cov(D_{\mathbf{i}}, D_{\mathbf{j}})\right|$. By Assumptions (P1) and (P2), we have

$$Var[D_{\mathbf{i}}] \leq C((\phi(\mathfrak{z}, r))^{(a+1)/a} + (\phi(\mathfrak{z}, r))^2);$$

therefore,

$$Q_{\mathbf{n}} = O\Big(p_{\mathbf{n}}^N \phi(\mathfrak{z}, r)\Big).$$

Concerning $\mathcal{R}_{\mathbf{n}}$, we introduce

$$S_1 = \{\mathbf{i}, \mathbf{j} \in I(1, \mathbf{n}, x, \mathbf{1}) : 0 < \|\mathbf{i} - \mathbf{j}\| \leq c_{\mathbf{n}}\},$$

$$S_2 = \{\mathbf{i}, \mathbf{j} \in I(1, \mathbf{n}, x, \mathbf{1}) : \|\mathbf{i} - \mathbf{j}\| > c_{\mathbf{n}}\},$$

where $c_{\mathbf{n}}$ is a real sequence that converges to $+\infty$. Split this sum over subsets in S_1 and S_2

$$\begin{aligned} \mathcal{R}_{\mathbf{n}} &= \Sigma_{(\mathbf{i},\mathbf{j}) \in S_1} \big|Cov\big(D_{\mathbf{i}}, D_{\mathbf{j}}\big)\big| + \Sigma_{(\mathbf{i},\mathbf{j}) \in S_2} \big|Cov\big(D_{\mathbf{i}}, D_{\mathbf{j}}\big)\big| \\ &= \mathcal{R}_{\mathbf{n}}^1 + \mathcal{R}_{\mathbf{n}}^2. \end{aligned}$$

First,

$$\begin{aligned} \mathcal{R}_{\mathbf{n}}^1 &= \Sigma_{(\mathbf{i},\mathbf{j}) \in S_1} \big|\mathbb{E}\big[F_{\mathbf{i}} F_{\mathbf{j}}\big] - \mathbb{E}[F_{\mathbf{i}}] \mathbb{E}\big[F_{\mathbf{j}}\big]\big| \\ &\leq C p_{\mathbf{n}}^N c_{\mathbf{n}}^N \phi(\mathfrak{z}, r) \Big((\phi(\mathfrak{z}, r))^{1/a} + \phi(\mathfrak{z}, r) \Big) \\ &\leq C p_{\mathbf{n}}^N c_{\mathbf{n}}^N (\phi(\mathfrak{z}, r))^{(a+1)/a}. \end{aligned}$$

On the other hand, we have

$$\mathcal{R}_{\mathbf{n}}^2 = \Sigma_{(\mathbf{i},\mathbf{j}) \in S_2} \big|Cov\big(D_{\mathbf{i}}, D_{\mathbf{j}}\big)\big|.$$

We deduce, from Lemma 4.1 in [8] that

$$\big|Cov\big(D_{\mathbf{i}}, D_{\mathbf{j}}\big)\big| \leq C\psi(\|\mathbf{i} - \mathbf{j}\|),$$

thus

$$\begin{aligned} \mathcal{R}_{\mathbf{n}}^2 &\leq C\Sigma_{(\mathbf{i},\mathbf{j}) \in S_2} \psi(\|\mathbf{i} - \mathbf{j}\|) \leq C p_{\mathbf{n}}^N \Sigma_{\mathbf{i}:\|\mathbf{i}\| \geq c_{\mathbf{n}}} \psi(\|\mathbf{i}\|) \\ &\leq C p_{\mathbf{n}}^N c_{\mathbf{n}}^{-Na} \Sigma_{\mathbf{i}:\|\mathbf{i}\| \geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \psi(\|\mathbf{i}\|). \end{aligned}$$

Let $c_{\mathbf{n}} = (\phi(\mathfrak{z}, r))^{-1/Na}$, then

$$\begin{aligned} \mathcal{R}_{\mathbf{n}}^2 &\leq C p_{\mathbf{n}}^N c_{\mathbf{n}}^{-Na} \Sigma_{\mathbf{i}:\|\mathbf{i}\| \geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \psi(\|\mathbf{i}\|) \\ &\leq C p_{\mathbf{n}}^N \phi(\mathfrak{z}, r) \Sigma_{\mathbf{i}:\|\mathbf{i}\| \geq c_{\mathbf{n}}} \|\mathbf{i}\|^{Na} \psi(\|\mathbf{i}\|). \end{aligned}$$

Because of (P2)

$$\mathcal{R}_{\mathbf{n}}^2 \leq C p_{\mathbf{n}}^N \phi(\mathfrak{z}, r).$$

Furthermore,

$$\mathcal{R}_{\mathbf{n}}^1 \leq C p_{\mathbf{n}}^N \phi(\mathfrak{z}, r).$$

Hence,

$$Var[X_1^*] = O\Big(p_{\mathbf{n}}^N \phi(\mathfrak{z}, r)\Big).$$

This last gives

$$\mathcal{B}_1 \leq \exp(-C(\eta_0) \ln \overline{\mathbf{n}}).$$

Consequently, a good choice of η_0 gives the claimed result of the lemma. Additionally,

$$\Sigma_{\mathbf{n}}\mathbb{P}\left(\left|\widehat{ES}_D(\mathfrak{z})\right| \leq 1/2\right) \leq \Sigma_{\mathbf{n}}\mathbb{P}\left(\left|\widehat{ES}_D(\mathfrak{z}) - \mathbb{E}\left[\widehat{ES}_D(\mathfrak{z})\right]\right| > 1/2\right) < \infty.$$

□

Lemma A2. *Under the supposition (P1)–(P2) and (P4)–(P5), we have*

$$\sup_{t\in[REXP_p(\mathfrak{z})-\delta, REXP_p(\mathfrak{z})+\delta]} \left|ES(t,\mathfrak{z}) - \mathbb{E}\left[\widehat{ES}_N(t,\mathfrak{z})\right]\right| = O\left(r^b\right).$$

Proof of Lemma A2. Writing

$$ES(t,x) - \mathbb{E}\left[\widehat{ES}_N(t,\mathfrak{z})\right] = \frac{1}{\mathbb{E}[F_{\mathbf{1}}(\mathfrak{z})]}\mathbb{E}\left[F_{\mathbf{1}}(\mathfrak{z})\mathbb{1}_{B(\mathfrak{z},r)}(\mathfrak{z}_1)(ES(t,x) - ES(t,A_{\mathbf{1}}))\right].$$

By (P2), we get

$$\mathbb{1}_{B(\mathfrak{z},r)}(A_{\mathbf{1}})|ES(t,x) - ES(t,A_{\mathbf{1}})| \leq Cr^b.$$

Thus,

$$\sup_{t\in[REXP_p(\mathfrak{z})-\delta, REXP_p(\mathfrak{z})+\delta]} |ES(t,x) - \mathbb{E}\left[\widehat{ES}_N(t,\mathfrak{z})\right]| \leq Cr^b,$$

which gives

$$\sup_{t\in[REXP_p(\mathfrak{z})-\delta, REXP_p(\mathfrak{z})+\delta]} |ES(t,x) - \mathbb{E}\left[\widehat{ES}_N(t,\mathfrak{z})\right]| = O(r^b)$$

□

Lemma A3. *Under the suppositions (P1)–(P5), we have*

$$\sup_{t\in[REXP_p(\mathfrak{z})-\delta, REXP_p(\mathfrak{z})+\delta]} \left|\widehat{ES}_N(t,\mathfrak{z}) - \mathbb{E}\left[\widehat{ES}_N(t,\mathfrak{z})\right]\right| = O(\left(\frac{\ln \overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)^{1/2}, \qquad a.co.$$

Proof of Lemma A3. Since $[REXP_p(\mathfrak{z}) - \delta,\ REXP_p(\mathfrak{z}) + \delta]$ then by the compactness feature we get

$$[REXP_p(\mathfrak{z}) - \delta,\ REXP_p(\mathfrak{z}) + \delta] \subset \bigcup_{j=1}^{l_{\mathbf{n}}}]B_j - d_{\mathbf{n}}, B_j + d_{\mathbf{n}}[\tag{A4}$$

for $d_{\mathbf{n}} = O\left(\frac{1}{\sqrt{\overline{\mathbf{n}}^b}}\right)$ and $l_{\mathbf{n}} = O\left(\sqrt{\overline{\mathbf{n}}^b}\right)$. The two functions $\mathbb{E}[\widehat{ES}_N(\cdot,\mathfrak{z})]$ and $\widehat{ES}_N(\cdot,\mathfrak{z})$ are increasing. Thus, for $1 \leq j \leq l_{\mathbf{n}}$,

$$\mathbb{E}\widehat{ES}_N((B_j - d_{\mathbf{n}},\mathfrak{z}) \leq \sup_{t\in]B_j-d_{\mathbf{n}},B_j+d_{\mathbf{n}}[} \mathbb{E}\widehat{ES}_N(t,\mathfrak{z}) \leq \mathbb{E}\widehat{ES}_N(B_j + d_{\mathbf{n}},\mathfrak{z})$$

$$\widehat{ES}_N(t,\mathfrak{z})B_j - d_{\mathbf{n}},\mathfrak{z}) \leq \sup_{t\in]B_j-d_{\mathbf{n}},B_j+d_{\mathbf{n}}[} \widehat{ES}_N(t,\mathfrak{z}) \leq \widehat{ES}_N(B_j + d_{\mathbf{n}},t). \tag{A5}$$

Now, by (P2)

$$\forall t_1, t_2 \in REXP_p(\mathfrak{z}) - \delta,\ REXP_p(\mathfrak{z}) + \delta,$$

we have

$$\left|\mathbb{E}\widehat{ES}_N(t_1,\mathfrak{z}) - \mathbb{E}\widehat{ES}_N(t_2,\mathfrak{z})\right| \leq C|t_1 - t_2|^b.$$

Hence,

$$\sup_{t\in[REXP_p(\mathfrak{z})-\delta,\, REXP_p(\mathfrak{z})+\delta]}\left|\widehat{ES}_N(t,\mathfrak{z})-\mathbb{E}\widehat{ES}_N(t,\mathfrak{z})\right|$$
$$\leq \max_{1\leq j\leq l_{\mathbf{n}}}\max_{z\in\{B_j-d_{\mathbf{n}},B_j+d_{\mathbf{n}}\}}\left|\widehat{ES}_N(z,\mathfrak{z})-\mathbb{E}\widehat{ES}_N(z,\mathfrak{z})\right|+Cd_{\mathbf{n}}^b.$$

Clearly,

$$d_{\mathbf{n}}^b=\overline{\mathbf{n}}^{-1/2}=o\left(\frac{\ln\overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)^{1/2}.$$

Therefore, it suffices that

$$\max_{1\leq j\leq l_{\mathbf{n}}}\max_{z\in\{B_j-d_{\mathbf{n}},B_j+d_{\mathbf{n}}\}}\left|\widehat{ES}_N(z,\mathfrak{z})-\mathbb{E}\widehat{ES}_N(z,\mathfrak{z})\right|=O\left(\frac{\ln\overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)^{1/2},\qquad a.co.$$

Then, $\forall\eta>0$,

$$\mathbb{P}\left(\max_{1\leq j\leq l_{\mathbf{n}}}\max_{z\in\{B_j-d_{\mathbf{n}},B_j+d_{\mathbf{n}}\}}\left|\widehat{ES}_N(z,\mathfrak{z})-\mathbb{E}\widehat{ES}_N(z,\mathfrak{z})\right|>\eta\sqrt{\frac{\ln\overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}}\right)$$
$$\leq 2l_{\mathbf{n}}\max_{1\leq j\leq l_{\mathbf{n}}}\max_{z\in\{B_j-d_{\mathbf{n}},B_j+d_{\mathbf{n}}\}}\mathbb{P}\left(\left|\widehat{ES}_N(z,\mathfrak{z})-\mathbb{E}\widehat{ES}_N(z,\mathfrak{z})\right|>\eta\sqrt{\frac{\ln\overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}}\right).$$

It remains to prove

$$\mathbb{P}\left(\left|\widehat{ES}_N(z,\mathfrak{z})-\mathbb{E}\widehat{ES}_N(z,\mathfrak{z})\right|>\eta\sqrt{\frac{\ln\overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}}\right).$$

Indeed,

$$\tilde{\mathbf{F}}_{\mathbf{i}}=\frac{1}{\mathbb{E}[F_{\mathbf{1}}]}\Big[F_{\mathbf{i}}B_{\mathbf{i}}\mathbb{1}_{\{B_{\mathbf{i}}\leq z\}}-\mathbb{E}\Big[F_{\mathbf{i}}B_{\mathbf{i}}\mathbb{1}_{\{B_{\mathbf{1}}\leq z\}}\Big]\Big].$$

We write $\forall\,\varepsilon>0$

$$\mathbb{P}\Big[|\widehat{ES}_N(z,\mathfrak{z})-\mathbb{E}\widehat{ES}_N(z,\mathfrak{z})|>\varepsilon\Big]=\mathbb{P}\left(\max_{z\in\mathcal{G}_{\mathbf{n}}}\left|\widehat{ES}_N(z,\mathfrak{z})-\mathbb{E}\Big[\widehat{ES}_N(z,\mathfrak{z})\Big]\right|>\varepsilon\right)$$
$$\leq\sum_{z\in\mathcal{G}_{\mathbf{n}}}\mathbb{P}\left(\left|\widehat{ES}_N(z,\mathfrak{z})-\mathbb{E}\Big[\widehat{ES}_N(z,\mathfrak{z})\Big]\right|>\varepsilon\right).\qquad\text{(A6)}$$

Since B is not necessarily bounded, we employ a truncation method by introducing

$$\widehat{ES}_N^*(\mathfrak{z},t)=\frac{1}{n\mathbb{E}[\mathbf{F}(h^{-1}d(\mathfrak{z},A_{\mathbf{1}}))]}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\mathbf{F}(r^{-1}d(\mathfrak{z},A_{\mathbf{i}}))B_{\mathbf{i}}^*$$

with $B^*=B\mathbb{1}_{(B<\gamma_{\mathbf{n}})}$ with $\gamma_{\mathbf{n}}=\overline{\mathbf{n}}^{a/p}$. Thus, the result is a consequence of

$$d_{\mathbf{n}}\max_{z\in\mathcal{G}_{\mathbf{n}}}\left|\mathbb{E}[\widehat{ES}_N^*(z,\mathfrak{z})]-\mathbb{E}[\widehat{ES}_N(z,\mathfrak{z})]\right|=O_{a.co.}\left(\frac{\ln\overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)^{1/2},\qquad\text{(A7)}$$

$$d_{\mathbf{n}}\max_{z\in\mathcal{G}_{\mathbf{n}}}\left|\widehat{ES}_N^*(z,\mathfrak{z})-\widehat{ES}_N(z,\mathfrak{z})\right|=O_{a.co.}\left(\frac{\ln\overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)^{1/2}\qquad\text{(A8)}$$

and

$$d_{\mathbf{n}}\max_{z\in\mathcal{G}_{\mathbf{n}}}\left|\widehat{ES}_N^*(z,\mathfrak{z})-\mathbb{E}[\widehat{ES}_N^*(z,\mathfrak{z})]\right|=O_{a.co.}\left(\frac{\ln\overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)^{1/2}.\qquad\text{(A9)}$$

For (A7) we write , $\forall\, z \in \mathcal{G}_{\mathbf{n}}$

$$\left|\mathbb{E}[\widehat{ES}^*_N(z,\mathfrak{z})] - \mathbb{E}[\widehat{ES}_N(z,\mathfrak{z})]\right| \le C\frac{1}{\phi(\mathfrak{z},r)}\mathbb{E}\left[|B|\mathbb{1}_{B\ge\gamma_{\mathbf{n}}\}}\mathbf{F}(r^{-1}d(\mathfrak{z},X))\right].$$

By the inequality of Holder , for α and β such that $\frac{1}{\alpha}+\frac{1}{\beta}=1$, and $\alpha=\frac{p}{2}$

$$\begin{aligned}
&\forall z \in \mathcal{G}_{\mathbf{n}}\\
\mathbb{E}\left[|B|\mathbb{1}_{\{B\ge\gamma_{\mathbf{n}}\}}\mathbf{F}(r^{-1}d(\mathfrak{z},A_1))\right] &\le \mathbb{E}^{1/\alpha}\left[|B^{\alpha}|\mathbb{1}_{\{B\ge\gamma_{\mathbf{n}}\}}\right]\mathbb{E}^{1/\beta}\left[\mathbf{F}^{\beta}(r^{-1}d(\mathfrak{z},A_1))\right]\\
&\le \gamma_{\mathbf{n}}^{-1}\mathbb{E}^{1/\alpha}\left[\left|B^{2\alpha}\right|\right]\mathbb{E}^{1/\beta}\left[\mathbf{F}^{\beta}(r^{-1}d(\mathfrak{z},A_1))\right]\\
&\le \gamma_{\mathbf{n}}^{-1}\mathbb{E}^{1/\alpha}[|B^{p}|]\mathbb{E}^{1/\beta}\left[\mathbf{F}^{\beta}(r^{-1}d(\mathfrak{z},A_1))\right]\\
&\le C\gamma_{\mathbf{n}}^{-1}\phi^{1/\beta}(\mathfrak{z},r).
\end{aligned}$$

Thus,

$$d_{\mathbf{n}}\max_{z\in\mathcal{G}_{\mathbf{n}}}\left|\widehat{ES}^*_N(z,\mathfrak{z}) - \mathbb{E}[\widehat{ES}^*_N(z,\mathfrak{z})]\right| \le \overline{\mathbf{n}}^{1/2-a/p}\phi^{(1-\beta)/\beta}.$$

Finally, (A7) is because $a>p$.

Now, for (A8) we use the Markov's inequality to show that $\forall z\in\mathcal{G}_{\mathbf{n}}$, $\forall\epsilon>0$

$$\begin{aligned}
\mathbb{P}\left(\left|\widehat{ES}^*_N(z,\mathfrak{z}) - \widehat{ES}_N(z,\mathfrak{z})\right| > \epsilon\right) &\le \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\mathbb{P}\left(B_i > n^{a/p}\right)\\
&\le \overline{\mathbf{n}}\mathbb{P}\left(B > n^{a/p}\right)\\
&\le \overline{\mathbf{n}}^{1-a}\mathbb{E}[B^p].
\end{aligned}$$

Choosing $\epsilon=\epsilon_0\left(\sqrt{\left(\frac{\ln\overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)}\right)$ and using $a>5/2$,

$$d_{\mathbf{n}}\max_{z\in\mathcal{G}_{\mathbf{n}}}\mathbb{P}\left(|\widehat{ES}_N(z,\mathfrak{z}) - \widehat{ES}^*_N(z,\mathfrak{z})| > \epsilon_0\left(\sqrt{\left(\frac{\ln\overline{\mathbf{n}}}{\overline{\mathbf{n}}\,\phi(\mathfrak{z},r)}\right)}\right)\right) \le \overline{\mathbf{n}}^{3/2-a} < C\overline{\mathbf{n}}^{-1-\nu}.$$

Now for (A9), define $z\in\mathcal{G}_{\mathbf{n}}$,

$$\mathcal{D}_{\mathbf{i}} = F_{\mathbf{i}}B^*_{\mathbf{i}} - \mathbb{E}[F_{\mathbf{1}}B^*_{\mathbf{i}}].$$

Therefore, $\forall\,\varepsilon>0$

$$\begin{aligned}
\mathbb{P}\left\{\left|\widehat{ES}^*_N(z,\mathfrak{z}) - \mathbb{E}\left[\widehat{ES}^*_N(z,\mathfrak{z})\right]\right| > \varepsilon\right\} &= \mathbb{P}\left\{\left|\frac{1}{\overline{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}]}\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\mathcal{D}_{\mathbf{i}}\right| > \varepsilon\right\}\\
&\le \mathbb{P}\left\{\left|\sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}\mathcal{D}_{\mathbf{i}}\right| > \varepsilon\overline{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}]\right\}.
\end{aligned}$$

Using the spatial blocks decomposition to write

$$\widehat{ES}^*_N(z,\mathfrak{z}) - \mathbb{E}\left[\widehat{ES}^*_N(z,\mathfrak{z})\right] = \frac{1}{\widehat{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}(x)]}\sum_{i=1}^{2^N}T(\mathbf{n},i), \tag{A10}$$

with

$$T(\mathbf{n},i) = \sum_{\mathbf{j}\in\mathcal{J}}\Lambda(i,\mathbf{n},\mathbf{j})$$

with

$$\text{where} \quad \mathcal{J} = \{0, \ldots, r_1 - 1\} \times \cdots \times \{0, \ldots, r_N - 1\}; \quad r_i = 2n_i p_{\mathbf{n}}^{-1},\, i = 1, \ldots, N.$$

and

$$\Lambda(1, \mathbf{n}, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}}+1 \\ k=1,\ldots,N}}^{2j_k p_{\mathbf{n}}+p_{\mathbf{n}}} \mathcal{D}_{\mathbf{i}},$$

$$\Lambda(2, \mathbf{n}, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}}+1 \\ k=1,\ldots,N-1}}^{2j_k p_{\mathbf{n}}+p_{\mathbf{n}}} \sum_{i_N = 2j_N p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{(j_N+1)p_{\mathbf{n}}} \mathcal{D}_{\mathbf{i}},$$

$$\Lambda(3, \mathbf{n}, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}}+1 \\ k=1,\ldots,N-2}}^{2j_k p_{\mathbf{n}}+p_{\mathbf{n}}} \sum_{i_{N-1} = 2j_{N-1} p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(j_{N-1}+1)p_{\mathbf{n}}} \sum_{i_N = 2j_N p_{\mathbf{n}}+1}^{2j_N p_{\mathbf{n}}+p_{\mathbf{n}}} \mathcal{D}_{\mathbf{i}},$$

$$\Lambda(4, \mathbf{n}, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}}+1 \\ k=1,\ldots,N-2}}^{2j_k p_{\mathbf{n}}} \sum_{i_{N-1} = 2j_{N-1} p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(j_{N-1}+1)p_{\mathbf{n}}} \sum_{i_N = 2j_N p_{\mathbf{n}}+p_{\mathbf{n}}+1}^{2(j_N+1)p_{\mathbf{n}}} \mathcal{D}_{\mathbf{i}},$$

$$\vdots$$

Finally

$$\Lambda(2^N, \mathbf{n}, \mathbf{j}) = \sum_{\substack{i_k = 2j_k p_{\mathbf{n}}+p_{\mathbf{n}}+1 \\ k=1,\ldots,N}}^{2(j_k+1)p_{\mathbf{n}}} \mathcal{D}_{\mathbf{i}}$$

Clearly, $T(\mathbf{n}, 1)$ is the sum of the random variables $\mathcal{D}_{\mathbf{i}}$ over big blocks, whereas the other terms $T(\mathbf{n}, i)$, $2 \leq i \leq 2^N$ are sums over small blocks.

Furthermore, from (A10), we get, for all $\eta > 0$,

$$\mathbb{P}\left(|\widehat{ES}^*_N(z, \mathfrak{z}) - \mathbb{E}\left[\widehat{ES}^*_N(z, \mathfrak{z})\right]| \geq \eta \right) \leq 2^N \max_{i=1,\ldots 2^N} \mathbb{P}(T(\mathbf{n}, i) \geq \eta \widehat{\mathbf{n}} \mathbb{E}[F_{\mathbf{1}}(x)]).$$

So, the required result is based on the evaluation of the quantities

$$\mathbb{P}(T(\mathbf{n}, i) \geq \eta \widehat{\mathbf{n}} \mathbb{E}[F_{\mathbf{1}}(x)]), \qquad \text{for all } i = 1, \ldots, 2^N.$$

For the sake of shortness, we treat only the case $i = 1$. The other case can be treated in the same manner. For the rest of the proof, we enumerate the $M = \prod_{k=1}^N r_k = 2^{-N} \widehat{\mathbf{n}} p_{\mathbf{n}}^{-N} \leq \widehat{\mathbf{n}} p_{\mathbf{n}}^{-N}$ random variables $\Lambda(1, \mathbf{n}, \mathbf{j})$; $\mathbf{j} \in \mathcal{J}$ in the arbitrary way $\mathcal{Z}_1, \ldots \mathcal{Z}_M$. Thus, for each $\mathcal{Z}_{\mathbf{j}}$, there exists a certain $\mathbf{j}$ in $\mathcal{J}$ such that

$$\mathcal{Z}_{\mathbf{j}} = \sum_{\mathbf{i} \in I(1, \mathbf{n}, \mathbf{j})} \mathcal{D}_{\mathbf{i}},$$

where $I(1, \mathbf{n}, \mathbf{j}) = \{\mathbf{i} : 2j_k p_{\mathbf{n}} + 1 \leq i_k \leq 2j_k p_{\mathbf{n}} + p_{\mathbf{n}} \quad ; k = 1, \ldots N\}$. Clearly the subsets $I(1, \mathbf{n}, \mathbf{j})$ contain $p_{\mathbf{n}}^N$ sites and are far apart by a distance of $p_{\mathbf{n}}$ at least. So, under (P4) and (P5),

$$\mathbf{F}(r_{\mathbf{n}}^{-1} d(x, A_{\mathbf{i}})) B^*_{\mathbf{i}} \leq C \gamma_{\mathbf{n}}.$$

So, according to the Lemma of [8] Carbon et al. (2007) we obtain M independent random variables $\mathcal{Z}^*_1, \ldots \mathcal{Z}^*_M$ having the same low as $\mathcal{Z}_{j=1,\ldots M}$ and such that

$$\sum_{j=1}^{r} \mathbb{E}|\mathcal{Z}_{\mathbf{j}} - \mathcal{Z}^*_{\mathbf{j}}| \leq 2C\gamma_{\mathbf{n}} M p_{\mathbf{n}}^N \phi(M-1) p_{\mathbf{n}}^N, p_{\mathbf{n}}^N) \phi(p_{\mathbf{n}}). \tag{A11}$$

Therefore,
$$\mathbb{P}(T(\mathbf{n},i) \geq \eta\widehat{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}(x)]) \leq \mathbf{B}_1(\mathbf{n}) + \mathbf{B}_2(\mathbf{n}),$$
where
$$\mathbf{B}_1(\mathbf{n}) = \mathbb{P}\left(\left|\sum_{j=1}^{M} \mathcal{Z}_{\mathbf{j}}^*\right| \geq \frac{M\eta\widehat{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}(x)]}{2M}\right)$$
$$\mathbf{B}_2(\mathbf{n}) = \mathbb{P}\left(\sum_{j=1}^{M} |\mathcal{Z}_{\mathbf{j}} - \mathcal{Z}_{\mathbf{j}}^*| \geq \frac{\eta\widehat{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}(x)]}{2}\right).$$
Concerning $\mathbf{B}_2(\mathbf{n})$, we write
$$\mathbf{B}_2(\mathbf{n}) \leq 2M\gamma_{\mathbf{n}} p_{\mathbf{n}}^N (\eta\widehat{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}(x)])^{-1}\phi((M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N)\psi(p_{\mathbf{n}}).$$
Now, since $\mathbb{E}[F_{\mathbf{1}}(x)] \leq C\phi(\mathfrak{z},r)$, $\widehat{\mathbf{n}} = 2^N M p_{\mathbf{n}}^N$ and $\phi((M-1)p_{\mathbf{n}}^N, p_{\mathbf{n}}^N) \leq p_{\mathbf{n}}^N$, we obtain for $\eta = \eta_0\sqrt{\dfrac{\ln\widehat{\mathbf{n}}}{\widehat{\mathbf{n}}\,\phi(\mathfrak{z},r)}}$
$$\mathbf{B}_2(\mathbf{n}) \leq \widehat{\mathbf{n}}\gamma_{\mathbf{n}} p_{\mathbf{n}}^N (\ln\widehat{\mathbf{n}})^{-1/2}(\widehat{\mathbf{n}}\phi(\mathfrak{z},r))^{-1/2}\psi(p_{\mathbf{n}}).$$
Therefore, for $p_{\mathbf{n}}$ $p_{\mathbf{n}} = C\left(\dfrac{\widehat{\mathbf{n}}\phi(\mathfrak{z},r)}{\ln\widehat{\mathbf{n}}\gamma_{\mathbf{n}}^2}\right)^{1/2N}$
$$\mathbf{B}_2(\mathbf{n}) \leq \widehat{\mathbf{n}}\,\psi(p_{\mathbf{n}}).$$
We conclude
$$\sum_{\mathbf{n}} \mathbf{B}_2(\mathbf{n}) < \infty.$$
Next, for $\mathbf{B}_1$,
$$\mathbf{B}_1(\mathbf{n}) \leq 2\exp\left(-\frac{(\eta\widehat{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}(x)])^2}{MVar\left[\mathcal{Z}_1^*\right] + C\eta\gamma_{\mathbf{n}} p_{\mathbf{n}}^N\widehat{\mathbf{n}}\mathbb{E}[F_{\mathbf{1}}(x)]}\right). \tag{A12}$$
Furthermore,
$$Var[\mathcal{Z}_1^*] = Var\left[\sum_{\mathbf{i}\in I(1,\mathbf{n},\mathbf{1})} \mathcal{D}_{\mathbf{i}}\right].$$
As $\mathbb{E}\left[B_{\mathbf{i}}^p|A_{\mathbf{i}}\right] < \infty$, for $p > 2$, then
$$\begin{aligned} Var\left[\mathcal{D}_{\mathbf{i}}^k\right] &\leq C\mathbb{E}\left[F_{\mathbf{i}}^2 B_{\mathbf{i}}^{*2}\right] \leq C\mathbb{E}\left[F_{\mathbf{i}}^2 B_{\mathbf{i}}^2\right] \\ &\leq C\mathbb{E}\left[F_{\mathbf{i}}^2\mathbb{E}\left[B_{\mathbf{i}}^2|A_{\mathbf{i}}\right]\right] \\ &\leq C\mathbb{E}\left[F_{\mathbf{i}}^2\right] \leq C\phi(\mathfrak{z},r), \end{aligned}$$
since $\mathbb{E}\left[|B_{\mathbf{i}}B_{\mathbf{j}}||A_{\mathbf{i}}A_{\mathbf{j}}\right] < \infty$ we get
$$\begin{aligned} \text{for all } \mathbf{i} \neq \mathbf{j}\ Cov(\mathcal{D}_{\mathbf{i}}, \mathcal{D}_{\mathbf{j}}) \quad &\leq C\mathbb{E}\left[F_{\mathbf{i}}|B_{\mathbf{i}}^*|F_{\mathbf{j}}|B_{\mathbf{j}}^*|\right] \\ &\leq C\mathbb{E}\left[F_{\mathbf{i}}F_{\mathbf{j}}|B_{\mathbf{i}}B_{\mathbf{j}}|\right] \\ &\leq C\mathbb{E}\left[F_{\mathbf{i}}F_{\mathbf{j}}\mathbb{E}\left[|B_{\mathbf{i}}B_{\mathbf{j}}||A_{\mathbf{i}}A_{\mathbf{j}}\right]\right] \\ &\leq C\mathbb{E}\left[F_{\mathbf{i}}F_{\mathbf{j}}\right] \leq C(\phi(\mathfrak{z},r))^{(a+1)/a}(h). \end{aligned}$$

Furthermore, as $\mathbb{E}\left[B_{\mathbf{i}}^p|A_{\mathbf{i}}\right] < \infty$

$$\begin{aligned}\text{for all } \mathbf{i} \neq \mathbf{j}\ Cov(\mathcal{D}_{\mathbf{i}}, \mathcal{D}_{\mathbf{j}}) &\leq \|\mathcal{D}_{\mathbf{i}}\|_p^2 \psi^{1-2/p}(\|i-j\|) \\ &\leq C\|F_{\mathbf{i}}B_{\mathbf{i}}^*\|_p^2 \psi^{1-2/p}(\|i-j\|) \\ &\leq C\|F_{\mathbf{i}}B_{\mathbf{i}}\|_p^2 \psi^{1-2/p}(\|i-j\|) \\ &\leq C\|F_{\mathbf{i}}\|_p^2 \psi^{1-2/p}(\|i-j\|) \\ &\leq C(\phi(\mathfrak{z}, r))^{2/p}(h)\psi^{1-2/p}(\|i-j\|)).\end{aligned}$$

Observe that

$$\sum_{\mathbf{i}\in I(1,\mathbf{n},\mathbf{1})} Var[\mathcal{D}_{\mathbf{i}}] = O\left(p_{\mathbf{n}}^N \phi(\mathfrak{z}, r)\right).$$

For a real sequence $d_{\mathbf{n}}$ tends to $+\infty$ we write

$$\begin{aligned}\sum_{\mathbf{i}\neq\mathbf{j}\in I(1,\mathbf{n},\mathbf{1})} \left|Cov(\mathcal{D}_{\mathbf{i}}, \mathcal{D}_{\mathbf{j}})\right| &\leq \sum_{\{\mathbf{i},\mathbf{j}\in I(1,\mathbf{n},\mathbf{1})\ \|\mathbf{i}-\mathbf{j}\|\leq d_{\mathbf{n}}\}} \left|Cov(\mathcal{D}_{\mathbf{i}}, \mathcal{D}_{\mathbf{j}})\right| \\ &+ \sum_{\{\mathbf{i},\mathbf{j}\in I(1,\mathbf{n},\mathbf{1})\ \|\mathbf{i}-\mathbf{j}\|> d_{\mathbf{n}}\}} \left|Cov(\mathcal{D}_{\mathbf{i}}, \mathcal{D}_{\mathbf{j}})\right| \\ &\leq Cp_{\mathbf{n}}^N \phi(\mathfrak{z}, r)\Big(d_{\mathbf{n}}^N(\phi(\mathfrak{z}, r))^{1/a} \\ &+ d_{\mathbf{n}}^{-Na}(\phi(\mathfrak{z}, r))^{2/p-1}(h)\textstyle\sum_{\mathbf{i}:\|\mathbf{i}\|\geq d_{\mathbf{n}}} \|\mathbf{i}\|^{Na}\psi^{1-2/p}(\|\mathbf{i}\|)\Big).\end{aligned}$$

Choosing $d_{\mathbf{n}} = (\phi(\mathfrak{z}, r))^{2/Np(a+1)-1/Na}$ to

$$\sum_{\mathbf{i}\neq\mathbf{j}\in I(1,\mathbf{n},\mathbf{1})} \left|Cov(\mathcal{D}_{\mathbf{i}}, \mathcal{D}_{\mathbf{j}})\right| \leq Cp_{\mathbf{n}}^N(\phi(\mathfrak{z}, r))$$

So,

$$Var\left[\sum_{\mathbf{i}\in I(1,\mathbf{n},\mathbf{1})} \mathcal{D}_{\mathbf{i}}\right] = O\left(p_{\mathbf{n}}^N(\phi(\mathfrak{z}, r))\right).$$

We replace $Var\left[\mathcal{Z}_1^*\right] = O\left(p_{\mathbf{n}}^N(\phi(\mathfrak{z}, r))\right)$ in (A12)

$$\mathbf{B}_1(\mathbf{n}) \leq \exp(-C(\eta_0)\ln\widehat{\mathbf{n}})$$

Finally, a good choice of η_0 gives

$$\sum_{\mathbf{n}} \mathbf{B}_1(\mathbf{n}) < \infty.$$

which completes the proof of the lemma. □

References

1. Cressie, N.A. *Statistics for Spatial Data*; Wiley: New York, NY, USA, 1993.
2. Diggle, P.; Ribeiro, P.J. *Model-Based Geostatistics*; Springer: New York, NY, USA, 2007.
3. Tran, L.T. Kernel density estimation on random fields. *J. Multivar. Anal.* **1990**, *34*, 37–53. [CrossRef]
4. Lu, Z.; Chen, X. Spatial kernel regression: Weak consistency. *Statist. Probab. Lett.* **2004**, *68*, 125–136. [CrossRef]
5. Biau, G.; Cadre, B. Nonparametric spatial prediction. *Stat. Inference Stoch. Process.* **2004**, *7*, 327–349. [CrossRef]
6. García-Soidán, P.H.; Febrero-Bande, M.; González-Manteiga, W. Nonparametric kernel estimation of an isotropic variogram. *J. Stat. Plan. Inference* **2004**, *121*, 65–92. [CrossRef]
7. Hallin, M.; Lu, Z.; Tran, L.T. Local linear spatial regression. *Ann. Stat.* **2004**, *32*, 2469–2500. [CrossRef]
8. Carbon, M.; Francq, C.; Tran, L.T. Kernel regression estimation for random fields. *J. Stat. Plan. Inference* **2007**, *137*, 778–798. [CrossRef]
9. Xu, R.; Wang, J. L^1-estimation for spatial nonparametric regression. *J. Nonparametr. Stat.* **2008**, *20*, 523–537. [CrossRef]

10. Li, J.; Tran, L.T. Nonparametric estimation of conditional expectation. *J. Stat. Plan. Inference* **2009**, *139*, 164–175. [CrossRef]
11. Dabo-Niang, S.; Yao, A.F. Kernel regression estimation for continuous spatial processes. *Math. Meth. Stat.* **2007**, *16*, 298–317. [CrossRef]
12. Laksaci, A.; Maref, F. Estimation non paramétrique de quantiles conditionnels pour des variables fonctionnelles spatialement dépendantes. *C. R. Math.* **2009**, *347*, 1075–1080. [CrossRef]
13. Mohammedi, M.; Bouzebda, S.; Laksaci, A. The consistency and asymptotic normality of the kernel type expectile regression estimator for functional data. *J. Multivar. Anal.* **2021**, *181*, 104673. [CrossRef]
14. Aneiros, G.; Cao, R.; Fraiman, R.; Genest, C.; Vieu, P. Recent advances in functional data analysis and high-dimensional statistics. *J. Multivar. Anal.* **2019**, *170*, 3–9. [CrossRef]
15. Almanjahie, I.M.; Bouzebda, S.; Kaid, Z.; Laksaci, A. The local linear functional kNN estimator of the conditional expectile: Uniform consistency in number of neighbors. *Metrika* **2024**, 1–29.. [CrossRef]
16. Litimein, O.; Laksaci, A.; Ait-Hennani, L.; Mechab, B.; Rachdi, M. Asymptotic normality of the local linear estimator of the functional expectile regression. *J. Multivar. Anal.* **2024**, *202*, 105281. [CrossRef]
17. Artzner, P.; Delbaen, F.; Eber, J.M.; Heath, D. Coherent measures of risk. *Math. Financ.* **1999**, *9*, 203–228. [CrossRef]
18. Righi, M.B.; Ceretta, P.S. A comparison of expected shortfall estimation models. *J. Econ. Bus.* **2015**, *78*, 14–47. [CrossRef]
19. Lazar, E.; Pan, J.; Wang, S. On the estimation of Value-at-Risk and Expected Shortfall at extreme levels. *J. Commod. Mark.* **2024**, *34*, 100391. [CrossRef]
20. Moutanabbir, K.; Bouaddi, M. A new non-parametric estimation of the expected shortfall for dependent financial losses. *J. Stat. Plan. Inference* **2024**, *232*, 106151. [CrossRef]
21. Scaillet, O. Nonparametric estimation and sensitivity analysis of expected shortfall. *Math. Financ. Int. J. Math. Stat. Financ. Econ.* **2004**, *14*, 115–129. [CrossRef]
22. Ferraty, F.; Quintela-Del-Río, A. Conditional VAR and expected shortfall: A new functional approach. *Econom. Rev.* **2016**, *35*, 263–292. [CrossRef]
23. Ait-Hennani, L.; Kaid, Z.; Laksaci, A.; Rachdi, M. Nonparametric estimation of the expected shortfall regression for quasi-associated functional data. *Mathematics* **2022**, *10*, 4508. [CrossRef]
24. Waltrup, L.S.; Sobotka, F.; Kneib, T.; Kauermann, G. Expectile and quantile regression—David and Goliath? *Stat. Model.* **2015**, *15*, 433–456. [CrossRef]
25. Bellini, F.; Di Bernardino, E.D. Risk management with expectiles. *Eur. J. Financ.* **2017**, *23*, 487–506. [CrossRef]
26. Bellini, F.; Negri, I.; Pyatkova, M. Backtesting VaR and expectiles with realized scores. *Stat. Methods Appl.* **2019**, *28*, 119–142. [CrossRef]
27. Farooq, M.; Steinwart, I. Learning rates for kernel-based expectile regression. *Mach. Learn.* **2019**, *108*, 203–227. [CrossRef]
28. Efron, B. Regression percentiles using asymmetric squared error loss. *Stat. Sin.* **1991**, *1*, 93–125.
29. Sobotka, F.; Kneib, T. Geoadditive expectile regression. *Comput. Stat. Data Anal.* **2012**, *56*, 755–767. [CrossRef]
30. Jiang, C.; Jiang, M.; Xu, Q.; Huang, X. Expectile regression neural network model with applications. *Neurocomputing* **2017**, *247*, 73–86. [CrossRef]
31. Daouia, A.; Girard, S.; Stupfler, G. Estimation of tail risk based on extreme expectiles. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2018**, *80*, 263–292. [CrossRef]
32. Maume-Deschamps, V.; Rullière, D.; Said, K. Multivariate extensions of expectiles risk measures. *Depend. Model.* **2017**, *5*, 20–44. [CrossRef]
33. Maume-Deschamps, V.; Rullière, D.; Said, K. Asymptotics multivariate expectiles. *arXiv* **2018**, arXiv:1704.07152v2.
34. Girard, S.; Stupfler, G.; Usseglio-Carleve, A. Functional estimation of extreme conditional expectiles. *Econom. Stat.* **2022**, *21*, 131–158. [CrossRef]
35. Goia, A.; Vieu, P. An introduction to recent advances in high/infinite dimensional statistics. *J. Multivar. Anal.* **2016**, *170*, 1–6. [CrossRef]
36. Yu, D.; Pietrosanu, M.; Mizera, I.; Jiang, B.; Kong, L.; Tu, W. Functional Linear Partial Quantile Regression with Guaranteed Convergence for Neuroimaging Data Analysis. *Stat. Biosci.* **2024**, 1–17. [CrossRef]
37. Di Bernardino, E.; Laloe, T.; Pakzad, C. Estimation of extreme multivariate expectiles with functional covariates. *J. Multivar. Anal.* **2024**, *202*, 105292. [CrossRef]
38. Guyon, X. Estimation d'un champ par pseudo-vraisemblance conditionnelle: Etude asymptotique et application au cas Markovien. In Proceedings of the Sixth Franco-Belgian Meeting of Statisticians, Bruxelles, Belguim, 14–15 November 1987.
39. Ferraty, F.; Vieu, P. *Nonparametric Functional Data Analysis: Theory and Practice*; Springer Series in Statistics; Springer: New York, NY, USA, 2006.
40. Li, W.V.; Shao, Q.M. Gaussian processes: Inequalities, small ball probabilities and applications. *Hanbook Stat.* **2001**, *19*, 533–597.
41. Liebscher, E. Estimation of the density and the regression function under mixing conditions. *Stat. Decis.* **2001**, *19*, 9–26. [CrossRef]
42. Rachdi, M.; Laksaci, A.; Al-Kandari, N.M. Expectile regression for spatial functional data analysis (sFDA). *Metrika* **2022**, *85*, 627–655. [CrossRef]
43. Hallin, M.; Lu, Z.; Yu, K. Local linear spatial quantile regression. *Bernoulli* **2009**, *15*, 659–686. [CrossRef]

44. Bayer, S.; Dimitriadis, T. Regression-Based Expected Shortfall Backtesting. *J. Financ. Econom.* **2022**, *20*, 437–471. [CrossRef]
45. Hassani, H.; Silva, E.S. A Kolmogorov-Smirnov based test for comparing the predictive accuracy of two sets of forecasts. *Econometrics* **2015**, *3*, 590–609. [CrossRef]