Science Journal of Applied Mathematics and Statistics
2026; 14(1): 27-36
http://www.sciencepublishinggroup.com/j/sjams
doi: 10.11648/j.sjams.20261401.14

Encompassing Test for Non-nested Linear and Nearest Neighbor Regressions Under Mixing Conditions

Patrick Rakotomarolahy*

LaMAF - Laboratory of Mathematics and their Applications of Fianarantsoa, University of Fianarantsoa, Fianarantsoa, Madagascar

Email address:
rakotopapa@gmail.com (Patrick Rakotomarolahy)

To cite this article:
Patrick Rakotomarolahy (2026). Encompassing Test for Non-nested Linear and Nearest Neighbor Regressions Under Mixing Conditions. *Science Journal of Applied Mathematics and Statistics, 14*(1), 27-36. https://doi.org/10.11648/j.sjams.20261401.14

Received: 24 November 2025 ; **Accepted:** 18 December 2025 ; **Published:** 16 January 2026

Abstract: Encompassing tests compare two non-nested competing models to determine whether one model contains all the relevant information captured by the other. They were developed based on the assumption of independence, particularly in cases involving parametric and non-parametric regression methods. However, this assumption is too restrictive since econometric dynamic models and time series rarely exhibit perfect independence. One challenge would be extending results for i.i.d. variables to dependent processes. We are interested in the ϕ-mixing dependence measure because of its properties, such as the fast decay rate. Consequently, the Central Limit Theorem (CLT) was previously obtained under significantly weaker mild conditions than those required for other mixing notions. This leads to better asymptotic behavior propreties for test statistics, despite practical constraints of the ϕ-mixing condition compared to weaker dependence measures. We examine the encompassing test for linear parametric and nonparametric nearest neighbor regression methods for ϕ-mixing processes. We establish the asymptotic normality of the encompassing statistics associated with the encompassing hypotheses. Our results are comparable to those obtained using popular methods in the literature, such as kernel regression. We achieve convergence rates of order $k^{-1/2}$ for the tests related to the nonparametric as rival model, and $n^{-1/2}$ for the parametric as rival model. These rates are the same as in the i.i.d. case. The asymptotic variances of the tests are well-defined due to the fast decay of the ϕ-mixing coefficients. Unlike many tests with nonparametric regression methods, ours do not depend on any density.

Keywords: Encompassing Test, Functional Parameter, ϕ-mixing Condition, Nearest Neighbor Regression, Asymptotic Normality

1. Introduction

Model selection is a challenging step in statistical modelling. Modelling any given dataset requires the associated data-generating process (DGP) to be characterised. As the DGP is unknown, we are faced with several admissible competing models. Model selection involves selecting a model that mimics the unknown DGP from a set of admissible models, according to a criterion. The model that makes the criterion optimal is chosen. Various model selection criteria exist in the literature for fully parametric admissible models, such as the Wald test, the likelihood ratio test, the Lagrange multiplier test and the information criteria. Some model selection procedures often rely on asymptotic theory when possible [44, 46, 49], or on bootstrap/resampling methods when asymptotic approximations are unreliable [8, 21, 27, 30]. The other case, where admissible models contain parametric and non-parametric specifications simultaneously, seems underdeveloped. An encompassing test appears to be helpful in the latter situation, where an encompassing model is intended to account for the salient features of the encompassed model. Therefore, an encompassing test can detect redundant models among the admissible models.

Encompassing tests are based on two principles: the encompassing model should be able to explain the predictions and predict mispecifications of the encompassed model [24]. Various considerations and developments of encompassing tests have been proposed [13, 14, 26, 34]. For an overview of the encompassing test concept [2, 35]. Applications of encompassing tests can be found in the GETS modelling procedure for model selection [23, 28].

Reference [1] developed encompassing tests for the linear parametric and the kernel nonparametric regression approaches. They showed that the statistics of these tests are asymptotically normally distributed, for independent and identically distributed variables (i.i.d.). Recently, an extension of this work can be found in [31, 42, 43] where [31, 43] focused on the situation with kernel regression. Additionally, reference [42] carried out an extension to nearest neighbour nonparametric regression methods.

We extend the results on the asymptotic behavior of encompassing tests under i.i.d. hypothesis in [1, 42]. We relax the i.i.d. hypothesis by considering stationary random processes with dependence structures. We establish the asymptotic normality of encompassing statistics associated to encompassing hypotheses for linear parametric and nonparametric nearest neighbour regression methods under ϕ-mixing dependence structure. This can be viewed as a generalisation of the encompassing test to dependent processes. Motivation on enhacing encompassing test for nearest neighbour regression may come from its consideration in the literature for dependent datasets or time series [17, 18, 37, 40]. Another motivation is that strong dependence yields faster convergence rates for test statistics in nonparametric regression. This opens up the development of encompassing tests for weaker dependence frameworks, allowing verifiable dependence conditions in nonlinear time series, Markov chains and econometric dynamic models [9–11, 33, 48]. These frameworks enable nonparametric inference in high-dimensional or dependent data.

The paper is organized as follows. We introduce the encompassing test for function parameters and the mixing condition for dependent processes in section 2. In section 3, we provide our results on the asymptotic normality of the encompassing test associated to linear parametric modelling and nonparametric nearest neighbor regression methods. Last we conclude.

2. Encompassing Test

This section introduces the encompassing hypothesis, the mixing dependence measures and assumptions.

2.1. Encompassing Hypothesis

We introduce the encompassing test and then build the corresponding encompassing hypothesis. So, given two regression models $\mathcal{M}_1$ and $\mathcal{M}_2$, we are interested in knowing if the model $\mathcal{M}_1$ can account the result of model $\mathcal{M}_2$. We want to know if $\mathcal{M}_1$ encompasses $\mathcal{M}_2$ or in a short notation $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$. Testing such a hypothesis will be done using the notion of encompassing test.

Generally speaking, model $\mathcal{M}_1$ encompasses model $\mathcal{M}_2$, if the parameter $\theta_{\mathcal{M}_2}$ of the latter model can be expressed in function of the parameter $\theta_{\mathcal{M}_1}$ of the former model. In other words, let $\Delta(\theta_{\mathcal{M}_1})$ be the pseudo true value of $\theta_{\mathcal{M}_2}$ on $\mathcal{M}_1$. In general, the pseudo-true value is defined as the plim of $\hat{\theta}_{\mathcal{M}_2}$ on $\mathcal{M}_1$ [1]. For more discussion on pseudo-true value associated with the KLIC[1] [16, 47]. The encompassing statistic is given by the difference between $\hat{\theta}_{\mathcal{M}_2}$ and $\Delta(\hat{\theta}_{\mathcal{M}_1})$ scaled by a coefficient a_n. Specification of the encompassing test will depend on the estimation of the regression method: parametric or nonparametric methods.

Let $S = (Y, X, Z)$ be a zero mean random process with valued in $\mathbb{R}\text{x}\mathbb{R}^d\text{x}\mathbb{R}^q$ where $d, q \in \mathbb{N}^*$. For $x \in \mathbb{R}^d$ and $z \in \mathbb{R}^q$, consider the two models $\mathcal{M}_1$ and $\mathcal{M}_2$ as the conditional expectations $m(x)$ and $g(z)$, respectively and are defined as follows:

$$\begin{array}{ll} \mathcal{M}_1 : & m(x) = E[Y|X = x] \quad and \\ \mathcal{M}_2 : & g(z) = E[Y|Z = z] \end{array} \tag{1}$$

Moreover, the general unrestricted model is given by $r(x, z) = E[Y|X = x, Z = z]$. We follow the encompassing test in [1].
We are interested in testing the hypothesis that $\mathcal{M}_1$ encompasses $\mathcal{M}_2$, and then introducing the null hypothesis:

$$\mathcal{H} : E[Y|X = x, Z = z] = E[Y|X = x]. \tag{2}$$

This null states that $\mathcal{M}_1$ is the owner model, and $\mathcal{M}_2$ will be served on validating this statement and is called the rival model. We test this hypothesis $\mathcal{H}$ through the following implicit encompassing hypothesis:

$$\mathcal{H}^* : E[E[Y|X = x]/Z = z] = E[Y|Z = z]. \tag{3}$$

The following homoskedasticity condition will be assumed all along this chapter:

$$Var[Y|X = x, Z = z] = \sigma^2. \tag{4}$$

Moreover, a necessary condition for the encompassing test relies on the errors of both models where the intended encompassing model $\mathcal{M}_1$ should have smaller standard error than the encompassed model $\mathcal{M}_2$.

In general, $\mathcal{M}_1$ or $\mathcal{M}_2$ can be estimated using nonparametric or parametric regression methods. We are interested on the asymptotic behavior of the encompassing statistic associated to the null hypothesis $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$. We can encounter the following four situations: $\mathcal{M}_1$ and $\mathcal{M}_2$ are both estimated parametrically, $\mathcal{M}_1$ and $\mathcal{M}_2$ are both estimated nonparametrically, $\mathcal{M}_1$ is estimated nonparametrically and $\mathcal{M}_2$ parametrically and $\mathcal{M}_1$ is estimated parametrically and $\mathcal{M}_2$ nonparametrically. We will consider the k-NN regression estimate for nonparametric methods and the linear regression for the parametric methods.

We now introduce the k-NN regression and various densities related to the process S, for which certain assumptions are required. Consider a sample $S_i = (Y_i, X_i, Z_i)$, $i = 1, ..., n$, which can be viewed as realization of the random process S. We suppose that S_i, $i = 1, ..., n$ has a joint density f. Moreover, $\varphi(.,.)$, $\varphi(. \mid .)$ and $\varphi(.)$ will denote the joint, the conditional and the marginal densities of the process (Y, Z),

[1] Kullback-Leiber information criterion

respectively. That is, for $y \in \mathbb{R}$ and $z \in \mathbb{R}^q$, $\varphi(y, z)$, $\varphi(y \mid z)$ and $\varphi(z)$ correspond to the density of the following processes (Y, Z) at point (y, z), $(Y \mid Z = z)$ at point y and Z at point z, respectively. Similarly, h will denote the joint, the conditional and the marginal densities of the process (Y, X), according to the argument that it takes.

The k-NN regression estimate g_n of the conditional mean g can be written as follows:

$$g_n(z) = \frac{\frac{1}{nR_n^q}\sum_{i=1}^n w(\frac{z-Z_i}{R_n})Y_i}{\frac{1}{nR_n^q}\sum_{i=1}^n w(\frac{z-Z_i}{R_n})} \tag{5}$$

where R_n will be defined as distance, according to the Euclidean norm in $\mathbb{R}^q$, from z to its k-th neighbors with k the number of neighbors, and $w(u)$ is a bounded, non negative function satisfying

$$\int w(u)du = 1 \quad and \quad w(u) = 0 \quad for \quad |u| \geq 1.$$

In sequel, the weighting function will be denoted by $w(z - Z_i)$ or simply by w_i, that is

$$w_i = \frac{\frac{1}{nR_n^q} w(\frac{z-Z_i}{R_n})}{\frac{1}{nR_n^q}\sum_{i=1}^n w(\frac{z-Z_i}{R_n})}.$$

We will specify the notion of dependence structure through the mixing conditions.

2.2. Mixing Dependence Measures

The concept of dependence may occur in different ways according to the structure of the process to be studied and could be ranged from the weakest to the strongest. Here, we consider some mixing conditions which measure the dependence structure on the process. They are the strong, the uniform and the regular mixings. Consider the process $\{S_i, i \in \mathbb{N}\}$ which is assumed to be stationary and denote by $\mathcal{F}_r^s = \sigma(S_\ell, r \leq \ell \leq s)$ the σ-algebra generated by $\{S_\ell, r \leq \ell \leq s\}$. We define the following mixing coefficients:

$$\alpha(n) = \underset{A\in\mathcal{F}_1^t, B\in\mathcal{F}_{n+t}^\infty}{Sup} |P(B \cap A) - P(A)P(B)|. \tag{6}$$

$$\beta(n) = \underset{B\in\mathcal{F}_{n+t}^\infty}{Sup} |P(B) - P(B \mid \mathcal{F}_1^t)|. \tag{7}$$

$$\phi(n) = \underset{A\in\mathcal{F}_1^t, B\in\mathcal{F}_{n+t}^\infty, P(A)\neq 0}{Sup} |P(B \mid A) - P(B)|. \tag{8}$$

Then the process $\{S_i, i \in \mathbb{N}\}$ is said to be α-mixing (or strong mixing) if $\alpha(n) \to 0$ as $n \to \infty$, β-mixing if $\beta(n) \to 0$ as $n \to \infty$ and ϕ-mixing (or uniform mixing) if $\phi(n) \to 0$ as $n \to \infty$. Moreover, such process is said to be geometrically strongly mixing if there exist $c_0 > 0$ and $\rho \in [0, 1[$ such that $\alpha(n) \leq c_0\rho^n$. For more development on the concepts of dependent processes [3, 4, 12, 29, 41, 45]. Example of dependent processes are the linear and the nonlinear parametric models, since processes like the ARMA models, related GARCH processes and Markov switching processes are known to be mixing [5, 7, 38]. We next specify the assumptions.

2.3. Assumptions

This section introduces the assumptions required to derive the asymptotic normality of the encompassing test. The first assumption is about the type of dependence measure.

Assumption 2.1. $(S_n)_n$ is ϕ-mixing process.

The next assumption relies on the joint, the conditional and the marginal density functions

Assumption 2.2. $g(z)$, $\varphi(y \mid z)$ and $\varphi(z)$ are p continuously differentiable. Moreover suppose that $\varphi(y \mid z)$ is bounded.

The last assumption is on the number of neighobrs and the weighting function of the nearest neighbor regression.

Assumption 2.3. The sequence $k = n^\beta$, $0 < \beta < \frac{2q}{2q+d}$ and such that $\sum_{t=1}^{k(n)} w_t = 1$ where w_t a weight function satisfying $0 < w_t < 1$ when $t \leq k(n)$ and $w_t = 0$ otherwise.

Under these assumptions, we will establish the asymptotic normality of the previous four encompassing tests under mixing conditions. We consider the first case where the two models $\mathcal{M}_1$ and $\mathcal{M}_2$ have parametric specifications.

3. Encompassing Test for the Fully Parametric Case

Encompassing test for parametric modelling has been developed a lot in the literature. We discuss briefly one parametric encompassing test where models $\mathcal{M}_1$ and $\mathcal{M}_2$ have linear parametric specification. In that case, the two models $\mathcal{M}_1$ and $\mathcal{M}_2$ are given by (9) and the nesting model is represented by the function r:

$$\begin{aligned} &m(x) = \beta' x \ \ with \ \ \beta = (E[XX'])^{-1}E[XY] \\ &g(z) = \gamma' z \ \ with \ \ \gamma = (E[ZZ'])^{-1}E[ZY] \\ &r(x,z) = \alpha' w \ \ with \ \ \alpha = E[WW']^{-1}E[WY] \ \ and \ \ W = (X, Z). \end{aligned} \tag{9}$$

We can get the estimates $\hat{\beta}$, $\hat{\gamma}$ and $\hat{\alpha}$ of the paramaters β, γ and α, respectively, using the sample $S_i = (Y_i, X_i, Z_i)$, $i = 1, ..., n$. Now, testing $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$ corresponds to the test of the null hypothesis $\mathcal{H}$ where the conditional mean is just the linear projection. Therefore, the encompassing statistic of the null $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$ can be written as follows.

$$\hat{\delta}_{\beta,\gamma} = \hat{\gamma} - \hat{\gamma_L}(\hat{\beta}), \tag{10}$$

where $\hat{\gamma_L}(\hat{\beta})$ is an estimate of the pseudo-true value $\gamma_L(\beta)$ associated with $\hat{\gamma}$ on $\mathcal{H}_1$. Remarking that the pseudo-true value is defined by $\gamma_L(\beta) = (E[ZZ'])^{-1}E[ZX']\beta$.

For development on this asymptotic behavior of the encompassing statistic, we refer to [15, 36]. For recent discussion on this encompassing test for fully parametric case [1].

Development of the parametric encompassing test goes beyond independent processes in the literature. We may mention the remark in [1] stating the obvious extension of the asymptotic results for independent processes to stationary ergodic processes in line with [52]. Moreover, encompassing test for dynamic stationary models and time series regressions have been discussed [16, 22, 25].

Next, we will study the completely nonparametric case.

4. Encompassing Test for the Fully Nonparametric Case

We now consider the case where the two models $\mathcal{M}_1$ and $\mathcal{M}_2$ defined in (1) are estimated using nonparametric techniques. To test the hypothesis " $\mathcal{M}_1$ encompasses $\mathcal{M}_2$ ", we build the corresponding encompassing statistic and establish asymptotic property of such statistic.

Using the sample $S_i = (Y_i, X_i, Z_i)$, $i = 1, ..., n$ and the associated functional estimates m_n and g_n of the unknown functions m and g in relation (1) respectively, we define the encompassing statistic as follows:

$$\hat{\delta}_{m,g}(z) = g_n(z) - \hat{G}(m_n)(z), \tag{11}$$

where $\hat{G}(m_n)$ is an estimate of the pseudo true value $G(m)$ associated with g_n on $\mathcal{H}$, which is defined by $G(m) = E[m \mid Z = z]$.

We establish asymptotic distribution of the encompassing statistic $\hat{\delta}_{m,g}$ in relation (11) when the conditional mean functions m and g defined in (1) are estimated using the k-NN regression method. We now provide asymptotic normality of the encompassing statistic of functional parameter for time series.

Theorem 4.1. Suppose that assumptions 2.1-2.3 hold, then under $\mathcal{H}$, we have:

$$\sqrt{k}\hat{\delta}_{m,g}(z) \to N(0, \gamma.\sigma^2) \quad in \quad distribution \quad as \quad n \to \infty. \tag{12}$$

where γ a positive constant which is equal to 1 when one considers uniform weights.

Proof of theorem 4.1. We have the following encompassing statistic when replacing the estimator of g_n and $\hat{G}(m_n)$ at a given point z:

$$\begin{aligned} \sqrt{k}\hat{\delta}_{m_n,g_n}(z) &= \sqrt{k}(\sum_{t=1}^{n} w(z - Z_t)Y_t) - \sqrt{k}(\sum_{t=1}^{n} w(z - Z_t)m_n(x_t)) \\ &= \sqrt{k}(\sum_{t=1}^{n} w(z - Z_t)(Y_t - m(x_t)) + \sqrt{k}(\sum_{t=1}^{n} w(z - Z_t)(m(x_t) - m_n(x_t)) \\ &= C + D \end{aligned} \tag{13}$$

where C is the first expression in RHS of the equality which is a k-NN regression of $\epsilon_t = Y_t - m(x_t)$ on Z_t scaled by the coefficient $\sqrt{k}$. Under $\mathcal{H}$ and with the assumptions of theorem 4.1, from Theorem 1 in [20], we have :

$$C \to N(0, \gamma.\sigma^2) \quad in \quad distribution \quad as \quad n \to \infty. \tag{14}$$

For the second expression D, we can bound it by taking its supremum with respect to x_t, so we get:

$$\begin{aligned} |D| &\leq Sup_{x_t}\sqrt{k}|m_n(x_t) - m(x_t)| \\ &\leq Sup_{x_t}\sqrt{k}|m_n(x_t) - E[m_n(x_t)]| + Sup_{x_t}\sqrt{k}|E[m_n(x_t)] - m(x_t)| \\ &= D1 + D2 \end{aligned} \tag{15}$$

Using the expression of the bias for ϕ-mixing processes in Lemma 1 [16], D_2 becomes:

$$D_2 = (Sup_{x_t}A(x_t))O((\frac{k}{n})^{\frac{(1-\beta)p}{d}})\sqrt{k} \tag{16}$$

which vanishes to zero when assumption 2.3 holds. $A(.)$ is a function which depends only on x_t and its expression can be found in [20].

Concerning D_1, we can obtain the following inequality:

$$\begin{aligned} D_1 &\leq Sup_{x_t}\sqrt{k}|m_n(x_t) - m_n(x)| + \sqrt{k}|m_n(x) - E[m_n(x)]| + Sup_{x_t}\sqrt{k}|E[m_n(x)] - E[m_n(x_t)]| \\ &= E1 + E2 + E3 \end{aligned} \tag{17}$$

If we denote by r_t the distance between x_t and its k_t^{th} neighbors, then the two nearest neighbor estimates $m_n(x_t)$ and $m_n(x)$ are given by:

$$m_n(x) = \sum_{s=1}^{n} w(x - X_s)Y_s \quad and \quad m_n(x_t) = \sum_{s=1}^{n} w(x_t - X_s)Y_s \tag{18}$$

where $w(x - X_s) = 0$ when $||x - X_s|| > r_n$ for all $s = 1, ..., n$ and $w(x_t - X_s) = 0$ when $||x_t - X_s|| > r_t$ or $s \geq t$. So, using relation (18), we have:

$$E_1 = Sup_{x_t} \sqrt{k} | \sum_{s=1}^{n} (w(x - X_s) - w(x_t - X_s))Y_s| \leq Sup_{x_t} \sum_{s=1}^{n} \frac{c_k}{\sqrt{k}} |Y_s| \tag{19}$$

where the last inequality comes from $Sup_{x_t}|(w(x - X_s) - w(x_t - X_s))| \leq \frac{c_k}{k}$ for some positive number c_k. If $Sup_{x_t} \sum_{s=1}^{\infty} ||Y_s|| < \infty$, E_1 is negligible. Else $Sup_{x_t} \sum_{s=1}^{\infty} |Y_s| = \infty$ which could be $Sup_{x_t} \sum_{s=1}^{n} |Y_s| \approx M\sqrt{log(n)}$ around infinity for a given positive real M, then $E_1 < M\sqrt{\frac{log(n)}{k}}$ which goes to zero as n goes to infinity under assumption 2.3.

Concerning E_2, for a given non zero real sequences a_{tn}, we have:

$$E_2 = \sqrt{k} | \sum_{t=1}^{n} (w(x - X_t)Y_t - Ew(x - X_t)Y_t)| = \sqrt{k} | \sum_{t=1}^{n} a_{tn}\psi_t| \tag{20}$$

where $\psi_t = \frac{1}{a_{tn}}(w(x - X_t)Y_t - Ew(x - X_t)Y_t)$. Applying Tchebyshev inequality on E_2, for $\epsilon > 0$ we have:

$$P(|E_2| > \epsilon) = P(\sqrt{k} | \sum_{t=1}^{n} a_{tn}\psi_t| > \epsilon) \leq \frac{k}{\epsilon^2} E(\sum_{t=1}^{n} a_{tn}\psi_t)^2 \tag{21}$$

We know that $E\psi_t = 0$ and ψ_t is a ϕ-mixing process, then from inequality of[50] relation (21) becomes: $P(E_1 > \epsilon) \leq \frac{k}{\epsilon^2} cA_n$ with c is a constant and $A_n = \sum_{t=1}^{n} a_{tn}^2$ where we choose a_{tn} according to $\frac{k}{\epsilon^2} cA_n \to 0$ as $n \to \infty$, then $E_2 \to 0$ in probability. For the last expression E_3, using relation (18), we obtain the following result:

$$E_3 = Sup_{x_t} \sqrt{k} |E[m_n(x)] - E[m_n(x_t)]| = Sup_{x_t} \sqrt{k} | \sum_{s=1}^{n} E[(w(x - X_s) - w(x_t - X_s))Y_s]| \tag{22}$$

Using iterated conditional expectation, we have:

$$\begin{aligned} E_3 &= Sup_{x_t} \sqrt{k} | \sum_{s=1}^{n} E[(w(x - X_s) - w(x_t - X_s))E[Y_s \mid X_s]]| \\ &= Sup_{x_t} \sqrt{k} | \sum_{s=1}^{n} E[(w(x - X_s) - w(x_t - X_s))m(X_s)]| \end{aligned} \tag{23}$$

One can write $m(X_s) = m(X_s) - m(x) + m(x)$ and then (23) becomes:

$$\begin{aligned} E_3 &= Sup_{x_t} \sqrt{k} | \sum_{s=1}^{n} E[(w(x - X_s) - w(x_t - X_s))(m(X_s) - m(x) + m(x))]| \\ &= Sup_{x_t} \sqrt{k} | \sum_{s=1}^{n} E[(w(x - X_s) - w(x_t - X_s))(m(X_s) - m(x))]| \end{aligned} \tag{24}$$

When $m(.)$ satisfies the Lipschitz condition, then relation (24) gives the following inequality:

$$\begin{aligned} E_3 &\leq Sup_{x_t}(c\sqrt{k} \sum_{s=1}^{n} E|w(x - X_s) - w(x_t - X_s)|||X_s - x||) \\ &\leq cr\sqrt{k} Sup_{x_t}(\sum_{s=1}^{n} E|w(x - X_s) - w(x_t - X_s)|) \end{aligned} \tag{25}$$

then under assumption 2.3 the RHS of the inequality (25) vanishes since $Sup_{x_t}(\sum_{s=1}^{n} E|w(x-X_s)-w(x_t-X_s)|) < \infty$. So E_3 goes to zero as $n \to \infty$. Thus, the proof is established.

Next, we will consider the mixed situation where the owner model has parametric specification and the rival is from nonparametric method.

5. Encompassing Test for Parametric vs Nonparametric Case

In this section, we consider the case that model $\mathcal{M}_1$ is a linear parametric model and $\mathcal{M}_2$ is estimated by nonparametric techniques. Therefore, through out of this section, the hypothesis $\mathcal{H}$ will have linear parametric specification. The encompassing statistic associated to the null $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$ can be written as follows:

$$\hat{\delta}_{\beta,g}(z) = g_n(z) - \hat{G}_L(\hat{\beta})(z), \tag{26}$$

where $\hat{G}_L(\hat{\beta})$ is an estimate of the pseudo-true value $G_L(\beta)(z)$ associated with g_n on $\mathcal{H}$, which is defined by $G_L(\beta)(z) = \beta' E[X \mid Z = z]$.

In this section, we assume that we estimate the rival model $\mathcal{M}_2$ using k-NN regression method where the owner model $\mathcal{M}_1$ is still with linear parametric specification. The following theorem provides the asymptotic behavior of the encompassing statistic introduced in relation (26).

Theorem 5.1. Assume that the relation (4) and Assumptions 2.1-2.3 are satisfied.
Then under $\mathcal{H}$, we get:

$$\sqrt{k-1}\hat{\delta}_{\beta,g}(z) \to N(0,\Sigma) \quad in \quad distribution \quad as \quad n \to \infty. \tag{27}$$

where $\Sigma = \gamma\sigma^2$ with γ a positive constant which is equal to 1 when one considers uniform weights.

We now provide the proof of this theorem.

Proof of Theorem 5.1. We now prove the case that the owner model $\mathcal{M}_1$ is the linear regression parametric and the rival model $\mathcal{M}_2$ is the k-NN regression nonparametric. We write the encompassing statistic as follows:

$$\begin{aligned}
\sqrt{k-1}\hat{\delta}_{\beta,g}(z) &= \sqrt{k-1}(g_n(z) - \hat{G}_L(\hat{\beta})(z)) \\
&= \sqrt{k-1}(\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})Y_i - \sum_{i=1}^{n} \tilde{W}(\frac{z-Z_i}{\tilde{R}_n})\hat{\beta}' X_i) \\
&= \sqrt{k-1}\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})(Y_i - \beta' X_i) + \sqrt{k-1}\sum_{i=1}^{n} \tilde{W}(\frac{z-Z_i}{\tilde{R}_n})X_i'(\beta - \hat{\beta}) \\
&= N_1 + N_2.
\end{aligned} \tag{28}$$

We mention that this equality 28 holds since the difference $\sqrt{k-1}\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})\beta' X_i - \sqrt{k-1}\sum_{i=1}^{n} \tilde{W}(\frac{z-Z_i}{\tilde{R}_n})\beta' X_i$ is negligible.

For the first expression $N_1 = \sqrt{k-1}\sum_{i=1}^{n} W(\frac{z-Z_i}{R_n})\epsilon_i$, with $\epsilon_i = Y_i - \beta' X_i$. When Assumptions 2.1-2.3 in Theorem 5.1 hold, then using result in [20], we have:

$$N_1 \to N(0,\Sigma) \quad in \quad distribution \quad as \quad n \to \infty. \tag{29}$$

where $\Sigma = c.\sigma^2 \int w^2(u)du$.

For $N_2 = (\beta - \hat{\beta})'\sqrt{k-1}\sum_{i=1}^{n} \tilde{W}(\frac{z-Z_i}{\tilde{R}_n})X_i$, under Assumptions 2.1-2.3, we know that the estimate $\sqrt{n}(\beta - \hat{\beta})$ converges in distribution to a normal law Z with mean zero and variance. The remaining part of N_2, the other expression $\frac{\sqrt{k-1}}{\sqrt{n}}\sum_{i=1}^{n} \tilde{W}(\frac{z-Z_i}{\tilde{R}_n})X_i$ converges in distribution to zero. Thus, from Slutsky's theorem, N_2 tends to zero in distribution. This result is achived from the uniform convergence in [20]. Thus, the proof of theorem 5.1 is established.

We now consider the last case where the owner model $\mathcal{M}_1$ is a nonparametric method and the rival model $\mathcal{M}_2$ is a linear parametric model.

6. Encompassing Test for Nonparametric vs Parametric Case

In this section, we consider the owner model $\mathcal{M}_1$ to be estimated using a nonparametric method and the rival model $\mathcal{M}_2$ to be a linear parametric method. Therefore, the encompassing statistic associated to the null $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$ is given by:

$$\hat{\delta}_{m,\gamma} = \hat{\gamma} - \hat{\gamma}(m_n), \tag{30}$$

where $\hat{\gamma}(m_n)$ is an estimate of the pseudo-true value $\gamma(m)$ associated with $\hat{\gamma}$ on $\mathcal{H}$, which is defined by $\gamma(m) = (E[ZZ'])^{-1}E[Zm]$. We estimate the unknown conditional mean m associated to the model $\mathcal{M}_1$ using the nonparametric k-NN regression estimate.

We state in the following theorem the asymptotic normality of the encompassing statistic in relation (30).

Theorem 6.1. Assume that relation 4 is satisfied. Moreover, suppose that Assumption 2.1 holds and the k-NN estimate m_n and the process $(Y_n, X_n)_n$ satisfy assumptions 2.2-2.3.
Then under $\mathcal{H}$, we get:

$$\sqrt{n}\hat{\delta}_{m,\gamma} \to N(0,\Omega) \quad in \quad distribution \quad as \quad n \to \infty. \tag{31}$$

where $\Omega = plim_{n\to\infty}Var(\sqrt{n}\hat{\delta}_{m,\gamma})$, in particular for independent case $\Omega = \sigma^2(E[Z'Z])^{-1}$.

Proof of Theorem 6.1. We mention that the functional parameters m_n is from k-NN regression estimate. We rewrite the associated encompassing statistic as follows:

$$\begin{aligned}\sqrt{n}\hat{\delta}_{m,\gamma} &= \sqrt{n}(\hat{\gamma} - \hat{\gamma}(m_n)) \\ &= \sqrt{n}((\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}(\frac{1}{n}\sum_{i=1}^{n} Z_i Y_i) - (\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}(\frac{1}{n}\sum_{i=1}^{n} Z_i m_n(x_i))) \\ &= \sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}(\frac{1}{n}\sum_{i=1}^{n} Z_i(Y_i - m(x_i))) \\ &+ \sqrt{n}(\frac{1}{n}\sum_{i=1}^{n} Z_i Z_i)^{-1}(\frac{1}{n}\sum_{i=1}^{n} Z_i(m(x_i) - m_n(x_i))) \\ &= L_1 + L_2.\end{aligned} \quad (32)$$

where L_1 corresponds to the first expression in the RHS of the equality (32). It coincides to the linear regression of the error ϵ on Z, with $\epsilon_i = Y_i - m(x_i)$.

When the processes sastify assumptions in the last theorem, then the normality asymptotic of the first expression L_1 is achieved using the normality asymptotic result for dependent processes in [41]. Similarly, the second expression L_2 vanishes from the same asymptotic normality of linear process in[41]. This completes the proof.

7. Remarks

In this section, we discuss the convergence rates and asymptotic variances of previous results on the asymptotic normality of various encompassing tests as well as the ϕ-mixing dependence measure. If possible, we will also attempt to draw comparisons with related results in the literature.

The asymptotic behaviors of encompassing statistics are driven by the rival model whose features are intended to be accounted in the owner model. This preserves the essence of this specific encompassing test among the model selection based on residuals, which is an important consideration. We then have two categories of convergence rates according to the rival model. We achieve convergence rates of order $k^{-1/2}$ for the nonparametric rival model, and $n^{-1/2}$ for the parametric rival model. These rates are the same as in the i.i.d. case [42]. This makes our results comparable with those obtained from kernel regression. The latter is equal to$h_n n^{q-1/2}$ where h_n is the bandwidth [1]. The method with the faster rate of convergence depends on the value of bandwidth h_n.

The mixing conditions typically affect the variance of the statistics, especially in a finite sample. However, when working with ϕ-mixing processes, the asymptotic variances of the tests are well-defined due to the fast decay of their mixing coefficients. An important issue for our encompassing statistics with the k-NN regression is that the asymptotic variances do not depend on the density. This kind of dependency typically arises in encompassing tests with kernel regression [1, 43].

The strength of a phi-mixing sequence lies in its ability to produce faster convergence rates for test statistics in nonparametric regression [51]. However, the application of the ϕ-mixing is limited in practice. Extending it to weaker dependence frameworks beyond classical mixing would significantly broaden the applicability of nonparametric tests, allowing the verification of dependence conditions in nonlinear time series, Markov chains,and econometric dynamic models. Under these frameworks, many tests retain asymptotic validity or non-asymptotic performance guarantees [10, 33]. This would enable nonparametric inference to be used in high-dimensional statistical analysis, representing an important step forward.

8. Conclusion

As stated in [24] that the work of [1] is the starting treatment of encompassing tests to functional parameter based on nonparametric methods. We have extended that work first to nearest neighbor functional parameter estimate for dependent process.

To summarize our results, testing the encompassing hypothesis $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$ needs asymptotic distribution of the corresponding statistic due to the lack of exact distribution. Specification of the encompassing test will depend on the type of $\mathcal{M}_1$ and $\mathcal{M}_2$. Modelling of $\mathcal{M}_1$ can be either parametric or nonparametric. We face the same situation for $\mathcal{M}_2$. We summarize in the table below the corresponding encompassing statistics.

We have focused on the three encompassing statistics without the completely parametric regression models. We have established the asymptotic normality, using linear and nearest neighbor regression methods, $\mathcal{M}_1\mathcal{E}\mathcal{M}_2$, that is

$$a_n\hat{\hat{\delta}}_{\theta_1,\theta_2} \to N(0,\Sigma) \quad in \quad distribution \quad as \quad n \to \infty. \quad (33)$$

where the order of convergence a_n and the asymptotic variance $\Sigma = plim_{n\to\infty} Var(a_n\hat{\hat{\delta}}_{\theta_1,\theta_2})$ depend on the regression method used and on the dependence structure of the processes.

One major limitation of our results is their application to real data, due to the stronger dependence structure. It is extremely challenging to verify the mixing condition in real data.

We would like to mention following possible perspectives. It would be interesting to obtain similar results with weaker dependence conditions. Developing encompassing tests that incorporate at least one nonparametric method could be seen as new research directions in theory and practice. The implementation of findings in any statistical software could also accelerate this progression and its application to econometric modeeling.

ORCID

0000-0002-7897-5356 (Patrick Rakotomarolahy)

Abbreviations

CLT	Central Limit Theorem
DGP	Data Generating Process
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GETS	General To Specific modeling
KLIC	Kullback-Leibler Information Criterion
k-NN	k-Nearest Neighbor

Acknowledgments

The author would like to thank the Editor and the two anonymous referees for their valuable comments, which have greatly improved the paper.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Bontemps, C., J. P. Florens and J. F. Richard (2008), "Parametric and non-parametric encompassing procedures", Oxford Bulletin of Economics and Statistics 70, 751-780. https://doi.org/10.1111/j.1468-0084.2008.00529.x

[2] Bontemps, C. and G. E. Mizon (2008). "Encompassing: concepts and implementation", Oxford Bulletin of Economics and Statistics 70, 721-750. https://doi.org/10.1111/j.1468-0084.2008.00528.x

[3] Bosq, D. (1998), "Nonparametric Statistics for Stochastic Processes: Estimation and Prediction", Lecture Notes in Statistics, Springer-Verlag, 2nd eds, Berlin. https://doi.org/10.1007/978-1-4612-1718-3

[4] Bradley, R. C. (2005), "Basic properties of strong mixing conditions: a survey and some open questions", Probability Surveys 2, 107-144. https://doi.org/10.1214/154957805100000104

[5] Carrasco, M. and X. Chen (2002), "Mixing and moment properties of various GARCH and Stochastic Volatility models", Econometric Theory 18, 17-39. https://doi.org/10.1017/S0266466602181023

[6] Cheng, P. E. (1984), "Strong consistency of nearest neighbor regression function estimators", Journal of Multivariate Analysis 15, 63-72. https://doi.org/10.1016/0047-259X(84)90067-8

[7] Davydov Y. A. (1974), " Mixing conditions for Markov chains", Theory Probab. Appl. 18, 312-328. https://doi.org/10.1137/1118033

[8] De Bin, R., Janitza, S., Sauerbrei, W., Boulesteix, A. L. (2016). Subsampling versus bootstrapping in resampling-based model selection for multivariable regression. Biometrics, 72(1), 272-280. https://doi.org/10.1111/biom.12381

[9] Dedecker, J., and Prieur, C. (2005). New dependence coefficients. Examples and applications to statistics. Probability Theory and Related Fields, 132(2), 203-236. https://doi.org/10.1007/s00440-004-0394-3

[10] Dedecker, J., Merlevède, F., and Rio, E. (2024). Deviation inequalities for dependent sequences with applications to strong approximations. Stochastic Processes and their Applications, 174, 104377. https://doi.org/10.1016/j.spa.2024.104377

[11] Dehling, H. and Wendler, M. (2010). Central limit theorem and the bootstrap for U-statistics of strongly mixing data. Journal of Multivariate Analysis 101(1), 126-137. https://doi.org/10.1016/j.jmva.2009.06.002

[12] Doukhan, P. (1994), "Mixing: properties and examples", Lecture Notes in Statistics 85, Springer Verlag.

[13] Florens, J. P., D. F. Hendry and J. F. Richard (1996), "Encompassing and specificity", Econometric Theory 12, 620-656. https://doi.org/10.1017/S0266466600006964

[14] Gouriéroux, C. and A. Monfort (1995), "Testing, encompassing, and simulating dynamic econometric models", Econometric Theory 11, 195-228. https://doi.org/10.1017/S0266466600009142

[15] Gouriéroux, C., A. Monfort and A. Trognon (1983), "Testing nested or non-nested hypotheses", Journal of Econometrics 21, 83-115. https://doi.org/10.1016/0304-4076(83)90121-5

[16] Govaerts, B., D. F. Hendry and J. F. Richard (1994), "Encompassing in stationary linear dynamic models", Journal of Econometrics 63, 245-270. https://doi.org/10.1016/0304-4076(93)01567-6

[17] Guégan, D. and N. Huck (2005), "On the use of nearest neighbors in finance", Revue de Finance 26, 67-86. https://doi.org/10.3917/fina.262.0067

[18] Guégan, D. and P. Rakotomarolahy (2009), "The multivariate k-nearest neighbor model for dependent variables : one-sided estimation and forecasting", CES Working Paper, No. 2009.50.

[19] Guégan, D. and P. Rakotomarolahy (2010), "A short note on the nowcasting and the forecasting of Euro-area GDP using non-parametric techniques", Economics Bulletin 30, 508-518.

[20] Guégan, D. and P. Rakotomarolahy, (2010), " Alternative methods for forecasting GDP", Chapter 8 in Fredj Jawadi, William A. Barnett (ed.) Nonlinear Modeling of Economic and Financial Time-Series (International Symposia in Economic Theory and Econometrics, vol. 20), Emerald Group Publishing Limited, 161-185.

[21] Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. The Annals of Statistics, 1926-1947. https://doi.org/10.1214/aos/1176349403

[22] Hendry, D. F. (1995), "Dynamic Econometrics", Oxford University Press, Oxford. https://doi.org/10.1093/0198283164.001.0001

[23] Hendry, D. F. and J. A. Doornik (1994), "Modelling linear dynamic econometric systems", Scottish Journal of Political Economy 41, 1-33. https://doi.org/10.1111/j.1467-9485.1994.tb01107.x

[24] Hendry, D. F., M. Marcellino and G. E. Mizon (2008), "Encompassing", Special Issue: Oxford Bulletin of Economics and Statistics, Guest Editor Introduction.

[25] Hendry, D. F. and B. Nielsen (2007), "Econometric modeling: a Likelihood Approach", Princeton University Press, Princeton. https://doi.org/10.1515/9781400845651

[26] Hendry, D. F. and J. F. Richard (1989), "Recent developments in the theory of encompassing", in Cornet B. and Tulkens H. (eds), Contributions to Operations Research and Economics, The XXth Anniversary of CORE, 393-440, Cambridge.

[27] Horowitz, J. L. (2019). Bootstrap methods in econometrics. Annual Review of Economics, 11(1), 193-224. https://doi.org/10.1146/annurev-economics-080218-025651

[28] Hoover, K. D. and S. J. Perez (1999), "Data mining reconsidered: encompassing and the general-to-specific approach to specification search", Econometrics Journal 2, 167-191. https://doi.org/10.1111/1368-423X.00025

[29] Ibragimov, I. A. (1962), "Some limit theorems for stationary processes" Theory of Probability and Application 7, 349-382. https://doi.org/10.1137/1107036

[30] James, G., Witten, D., Hastie, T., Tibshirani, R., Taylor, J. (2023). Resampling methods. In An introduction to statistical learning: With applications in python (pp. 201-228). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-38747-0_5

[31] Lapenta E and P. Lavergne (2025), " Encompassing Tests for Nonparametric Regressions" Econometric Theory 41, 709-738. https://doi.org/10.1017/S0266466624000100

[32] Mack, Y. P. (1981), "Local properties of k-NN regression estimates", SIAM Journal on Algebraic and Discrete Methods 2, 311-323. https://doi.org/10.1137/0602035

[33] Merlevède F., Peligrad, M., and Utev, S. (2019). Functional CLT for martingale-like nonstationary dependent structures. Bernoulli, 25(4B), 3205-3225. https://doi.org/10.3150/18-BEJ1088

[34] Mizon, G. E. (1984), "The encompassing approach in econometrics", in D. F. Hendry and K. F. Wallis (eds), Econometrics and Quantitative Economics, 135-172.

[35] Mizon, G. E. (2008). "Encompassing", in Blume L. E. and Durlauf S. N. (eds), The New Palgrave Dictionary of Economics, 2nd edn. The New Palgrave Dictionary of Economics Online, Palgrave Macmillan. https://doi.org/10.1007/978-1-349-58802-2_469

[36] Mizon, G. E. and J. F. Richard (1986), "The encompassing principle and its application to non-nested hypothesis tests", Econometrica 54, 657-678. https://doi.org/10.2307/1911313

[37] Mizrach, B. (1992), "Multivariate Nearest-Neighbour Forecasts of EMS Exchange Rates", Journal of Applied Econometrics 7, Supplement: Special Issue on Nonlinear Dynamics and Econometrics (Dec., 1992), S151-S163. https://doi.org/10.1002/jae.3950070511

[38] Mokkadem, A. (1990), "Propriétés de mélange des modèles autoregressifs polynomiaux", Annales de l'Institut Henri Poincaré 26, 219-260.

[39] Mukerjee, H. (1993), "Nearest neighbor regression with heavy-tailed errors", Annals of Statistics 21, 681-693. https://doi.org/10.1214/aos/1176349144

[40] Nowman, B. and B. Saltoglu (2003), "Continuous time and nonparametric modelling of U. S. interest rate models", International Review of Financial Analysis 12, 25-34. https://doi.org/10.1016/S1057-5219(02)00123-0

[41] Peligrad, M. and S. A. Utev (1997), "Central limit theorem for Linear Processes", The Annals of Probability 25, 443-456. https://doi.org/10.1214/aop/1024404295

[42] Rakotomarolahy P. (2020), "Asymptotic behavior of encompassing test for independent processes: case of linear and nearest neighbor regressions", Cogent Mathematics and Statistics 7, 1805092. https://doi.org/10.1080/25742558.2020.1805092

[43] Rakotomarolahy P. (2023), 'Asymptotic normality of the encompassing test associated to the linear parametric modelling and the kernel method for alpha mixing processes", London Journal of Research In Science: Natural and Formal 23 (7), 11-21.

[44] Rivers, D. and Vuong, Q. (2002). Model selection tests for nonlinear dynamic models. The Econometrics Journal, 5(1), 1-39. https://doi.org/10.1111/1368-423X.t01-1-00071

[45] Rosenblatt, M. (1956a), "A central limit theorem and a strong mixing condition", Proc. Natl. Acad. Sci. USA 42, 43-47. https://doi.org/10.1073/pnas.42.1.43

[46] Shao, J. (1997). An asymptotic theory for linear model selection. Statistica sinica, 221-242.

[47] Sawa, T. (1978), "Information criteria for discriminating among alternative regression models", Econometrica 46, 1273-1292. https://doi.org/10.2307/1913828

[48] Vieu, P. (1994), "On variable selection in nonparametric regression", Computation Statistics and Data Analysis 17, 575-594. https://doi.org/10.1016/0167-9473(94)90149-X

[49] Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica: journal of the Econometric Society, 307-333. https://doi.org/10.2307/1912557

[50] Yoshihara, K. (1978). "Probability inequalities for sums of absolutely regular processes and their applications", Zeitschrift fuer Wahrscheinlichkeitstheorie und Verwandte Gebiete 43, 319-329. https://doi.org/10.1007/BF00534765

[51] Yu, B. (1994). Rates of convergence for empirical processes of stationary mixing sequences. The Annals of Probability 22(4), 94-116. https://doi.org/10.1214/aop/1176988849

[52] White, H. (1990), "A consistent model selection", in Granger C. W. J. (ed), Modelling Economic Series, Clarendon Press, Oxford, 369-383.