

Un modelo de regresión con ruido fraccionario coloreado



Silfrido Gómez Polanco

Supervisor: Prof. Soledad Torres

Advisor: Prof. Lisandro Fermin

Instituto de Estadística
Universidad de Valparaíso

This dissertation is submitted for the degree of
Doctor en Estadística

Agradecimientos

Al espacio, al universo y Dios, por permitir que me mantenga cada día en mi lugar y tener siempre una nueva oportunidad de vida. Y por otro lado, pero muy especialmente a mi madre, por ser su mejor versión al momento de concebirme y criarme, pues siempre me brinda la dicha de compartir a su lado incondicionalmente.

Resumen

Versión Español

En la presente investigación se obtuvieron dos resultados importantes, el primero de ellos y el principal, fue la estimación de un modelo de regresión espacio-temporal ponderado, dirigido por un ruido fraccionario coloreado. Es importante resaltar, que para dicho modelo también se logró obtener una forma explícita de la función de covarianza, la cual expresa de mejor manera la información conjunta observada. El segundo resultado, el cual es clave para trabajos futuros, fue la obtención de la forma explícita de las soluciones de la ecuación del calor en forma discretizada. Estas, obtenidas a partir de una discretización sobre una base finita de auto-funciones del operador de segundo orden para dicha ecuación del calor. Es importante mencionar, que estas soluciones son un modelo autoregresivo ponderado y que se identifican con el modelo autorregresivo geográfico y temporalmente ponderado GTWAR.

Cabe destacar, que para el primer resultado, fue necesario probar la consistencia del estimador de mínimos cuadrados ponderados para el modelo propuesto, a modo de evidenciar el comportamiento asintótico este. De esto, resultó un gran teorema, el cual prueba bajo supuesto de fuerte interacción entre la covariables, la convergencia en probabilidad y la convergencia casi segura del estimador mencionado.

Finalmente, como segundo resultado, el cual abre trabajos futuros como bien ya se dijo, logramos escribir la forma discretizada de las soluciones para la ecuación del calor, permitiendo este hecho reconocer a la solución encontrada como un modelo autoregresivo espacial y temporalmente ponderado. Así, quedaría por continuar la identificación de los coeficientes entre el modelo GTWAR y la consistencia del estimador de mínimos cuadrado ponderados para un GTWAR con ruido fraccionario coloreado.

English Version

In the present investigation, two important results were obtained, the first and main one was the estimation of a weighted spatio-temporal regression model, directed by colored

fractional noise. It is important to highlight that for said model it was also possible to obtain an explicit form of the covariance function, which better expresses the joint information observed. The second result, which is key for future work, was obtaining the specific form of the solutions of the heat equation in discretized form. These, obtained from a discretization on a finite basis of eigenfunctions of the second order operator for said heat equation. It is important to mention that these solutions are a weighted autoregressive model and that they are identified with the geographical and temporally weighted GTWAR autoregressive model.

It should be noted that for the first result, it was necessary to test the consistency of the weighted least squares estimator for the proposed model, in order to demonstrate its asymptotic behavior. From this, a great theorem resulted, which proves, under the assumption of strong interaction between the covariates, the convergence in probability and the almost certain convergence of the aforementioned estimator.

Finally, as a second result, which opens up future work as already stated, we managed to write the discretized form of the solutions for the heat equation, allowing this fact to recognize the solution found as a spatially and temporally weighted autoregressive model. Thus, it would remain to continue the identification of the coefficients between the GTWAR model and the consistency of the weighted least squares estimator for a GTWAR with colored fractional noise.

Índice general

Índice de figuras	IX
Índice de cuadros	XI
1. Introducción	1
1.1. Modelo de Regresión Ponderado Temporal y Geográficamente (GTWR) . . .	2
1.2. Modelo Auto-Regresivo Ponderado Temporal y Geográficamente - (GTWAR) 10	
1.3. Ecuaciones en derivadas parciales estocásticas (SPEDs)-dirigidas por un ruido coloreado fraccionario	13
1.4. Contribuciones	14
2. Modelo de regresión geográfica y temporalmente ponderada con ruido coloreado fraccionario	17
2.1. Modelo GTWR	18
2.2. Movimiento Browniano Fraccionario (mBf)	19
2.3. Ruido Coloreado Fraccionario	21
2.4. Covariables correlacionadas	25
2.5. Estimador de mínimos cuadrados ponderados	26
2.6. Consistencia del estimador	28
2.7. Simulaciones	42
2.7.1. Ruido coloreado con ($H_s = 0,40$ y $H_t = 0,65$)	43
2.7.2. Simulación: modelo ($Y = \beta_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$)	44
2.7.3. Estimación: modelo ($\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$)	47
2.7.4. Residuos: modelo ($\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$)	48
2.7.5. Simulación: modelo ($Y = \beta_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$)	50
2.7.6. Estimación: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$)	52
2.7.7. Residuos: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$)	53
2.7.8. Ruido coloreado con ($H_s = 0,40$ y $H_t = 0,90$)	56

2.7.9. Simulación: modelo ($Y = \beta_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)	57
2.7.10. Estimación: modelo ($\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)	59
2.7.11. Residuos: modelo ($\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)	60
2.7.12. Simulación: modelo ($Y = \beta_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)	63
2.7.13. Estimación: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)	65
2.7.14. Residuos: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)	66
3. Trabajo futuros: (Estimación de parámetros en SPEDs asociadas a un modelo GTWAR)	69
3.1. SPDEs con ruido fraccionario coloreado	70
3.2. Discretización de soluciones	72
Apéndice A.	75
Bibliografía	77

Índice de figuras

1.1. Esquema del modelo de regresión geográfico y temporal	3
1.2. Un ejemplo de ancho de banda espaciotemporal de decaimiento en el tiempo.	5
1.3. Matriz W_i de los pesos individuales $w_{ij,s,t}^t$	6
2.1. Ruido coloreado espacial	43
2.2. Modelo $Y = \beta_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	44
2.3. $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	47
2.4. Residuos: modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	48
2.5. Histogramas de los residuos para el modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	48
2.6. Gráficos de QQ-plot: modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	49
2.7. (ECM): modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	49
2.8. Box-plot para $\hat{\beta}_1$	50
2.9. Modelo $Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	50
2.10. \hat{y} para β_2	53
2.11. Residuos: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$)	53
2.12. Histogramas de los residuos para el modelo $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	54
2.13. Gráficos de QQ-plot: modelo $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	54
2.14. (ECM): modelo $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	55
2.15. Box-plot para $\hat{\beta}_2$	55
2.16. Ruido coloreado espacial	56
2.17. Modelo $Y = \beta_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$	57
2.18. Estimación: modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$	60
2.19. Residuos: modelo ($\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$)	60
2.20. Histogramas de los Residuos: modelo ($\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$) . .	61
2.21. Gráficos Q-Q plot de los residuos para $\hat{\beta}_1$	61
2.22. (ECM): modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$	62
2.23. Box-plot para $\hat{\beta}_1$	62
2.24. Modelo $Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$	63

2.25. $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$	65
2.26. Residuos: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$)	66
2.27. Histogramas de los residuos para β_1	66
2.28. Gráficos Q-Q plot de los residuos para $\hat{\beta}_2$	67
2.29. Error cuadrático medio para $\hat{\beta}_2$	67
2.30. Box-plot $\hat{\beta}_2$	68

Índice de cuadros

2.1.	Resultados de la regresión global $Y = \beta_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	45
2.2.	Resultados de la regresión global $Y = \beta_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	46
2.3.	Resultados de la regresión global $Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	51
2.4.	Resultados de la regresión global $Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$	52
2.5.	Resultados de la regresión global $Y = \beta_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$	58
2.6.	Resultados de la regresión global $Y = \beta_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$	59
2.7.	Resultados de la regresión global $Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$	63
2.8.	Resultados de la regresión ponderada $Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$	64

Capítulo 1

Introducción

El tiempo y el espacio son dos dimensiones esenciales que delimitan las actividades humanas, los eventos sociales y en particular los procesos ambientales. Análisis y visualización de ambas dimensiones, tomando como enfoque el modelamiento espacio-temporal, ha sido durante mucho tiempo un importante foco de investigación en el campo de la ciencia de la información geográfica (GIScience) Cressie (1993). Un ejemplo de ello es el marco teórico de la geografía del tiempo propuesto por Hägerstrand (1970) Ellegård and Svedin (2012), en donde la disponibilidad y abundancia de datos espacio-temporales, como en las trayectorias recopiladas por los sistemas de posicionamiento global (GPS) y las imágenes instantáneas de teledetección, motivaron un interés particular en el modelamiento espacio-temporal. Algunos ejemplos de lo que se menciona incluyen la exploración de patrones espacio-temporales del comportamiento humano (Chen et al. 2011; Kwan 2000, 2004) Chen et al. (2011), modelado de la variación en precios de viviendas según su georeferenciación Fotheringham et al. (2002), actividades delictivas (Brunsdon, Corcoran y Higgs 2007; Nakaya y Yano 2010) Cheng and Williams (2012), brotes de enfermedades (Takahashi et al. 2008)Takahashi et al. (2008), como también métodos para analizar y visualizar datos en espacio y tiempo (Andrienko et al. 2010; Demšar y Virrantaus 2010; Rey y Janikas 2010)Fotheringham et al. (2015b).

Ahora bien, más allá de los particulares problemas que pueda encerrar cada ámbito de estudio asociado a dichas áreas de investigación que se mencionan en el párrafo anterior, se encuentra el hecho de que en cada una de ellas se conoce cierta información observada, la cual mediante su entorno y las características específicas en el tiempo y el espacio que presentan, permiten usar modelos lineales para describir la relación entre las variables que representen dicha información del entorno y lo que se pueda estar observando en las dimensiones tanto temporal como espacial. En ese sentido y dicho esto, haremos referencia al modelo de regresión geográfica y temporalmente ponderado o como sus siglas en inglés lo describen **(GTWR) (Geographically and Temporally Weighted Regression model)**.

1.1. Modelo de Regresión Ponderado Temporal y Geográficamente (GTWR)

El modelo GTWR es un modelo de regresión lineal ponderado espacial y temporalmente, su expresión tiene la siguiente forma:

$$y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^p \beta_k(u_i, v_i, t_i)x_{i,k} + \varepsilon_i, \quad i = 1, \dots, n \quad (1.1)$$

donde i es el índice de la i_{th} -ésima observación espacio-temporal, (u_i, v_i, t_i) denota sus coordenadas en espacio y tiempo, con y_i la variable dependiente, x_{ik} como k_{th} -ésima variable independiente, ε_i el término de error para la i_{th} -ésima observación, $\beta_0(u_i, v_i, t_i)$ representa el intercepto y $\beta_k(u_i, v_i, t_i)$ es un conjunto de valores de parámetros para la i_{th} -ésima observación espacio-temporal.

Este modelo es una extensión del modelo **GWR (Geographically Weighted Regression)** ver Fotheringham et al. (2002). Es por ello, que antes de avanzar con los detalles que permiten conocer su estructura y uso, revisaremos cómo se contruye el GWTR conectando y analizando las ideas planteadas por los autores para la extensión a un modelo más complejo. Esto, lógicamente nos conducirá a la explicación de los elementos que constituyen al GTWR enunciado inicialmente.

Iniciemos con la definición del GWR basándonos en los trabajos de Fotheringham et al. (2002), con la intención de mostrar los cambios más importantes en la estructura de dicho modelo, que permitieron su extensión al GTWR.

El modelo GWR se representa mediante la siguiente expresión:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i)x_{i,k} + \varepsilon_i, \quad i = 1, \dots, n \quad (1.2)$$

donde i es el índice de un punto espacial, con (u_i, v_i) denotando sus coordenadas. En consecuencia, $y_i, x_{ik}, \varepsilon_i$ son: variable dependiente, k -ésima variable independiente y término de error para la i -ésima observación, respectivamente. Por otra lado, $\beta_0(u_i, v_i)$ representa el intercepto y $\beta_k(u_i, v_i)$ es un conjunto de valores de parámetros en el punto i .

El principal atractivo de este modelo, es que los parámetros $\beta_k(u_i, v_i)$ pueden variar a través del espacio para medir relaciones espacialmente no estacionarias. Más aún, si se utiliza una representación matricial, los parámetros a estimar vía mínimos cuadrados en el modelo se les puede representar mediante la ecuación:

$$\hat{\beta}(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) y \quad (1.3)$$

donde W es una matriz diagonal con elementos que representan los pesos geográficos de cada observación del i -ésimo punto. Ahora bien, cuando las observaciones se recopilan a lo largo del tiempo y el espacio, en un conjunto de locaciones $S_t = [s_1, s_2, s_3, \dots, s_{n_t}]$ donde n_t es el número de ubicaciones en las que se observan los datos en el período de tiempo t , el modelo GWR en la ecuación 1.1 aún se puede usar para derivar estimaciones locales, pero en este caso incorporando los datos medidos en períodos de tiempo anteriores $t - 1, t - 2, \dots, t - q$, siendo q el número de lapsos de tiempo adicional al del mismo período de tiempo. Así, la estimación de los parámetros todavía puede obtenerse mediante la ecuación 1.3. La diferencia, sin embargo, radica en que para el GTWR la matriz de peso W , se construye de una manera diferente, pues captura los efectos espaciales y temporales de las observaciones cercanas tanto en el espacio como en el tiempo. Ver Fotheringham et al. (2015a).

Una forma ordenada de comprender el funcionamiento del GTWR se puede ver en el siguiente esquema:

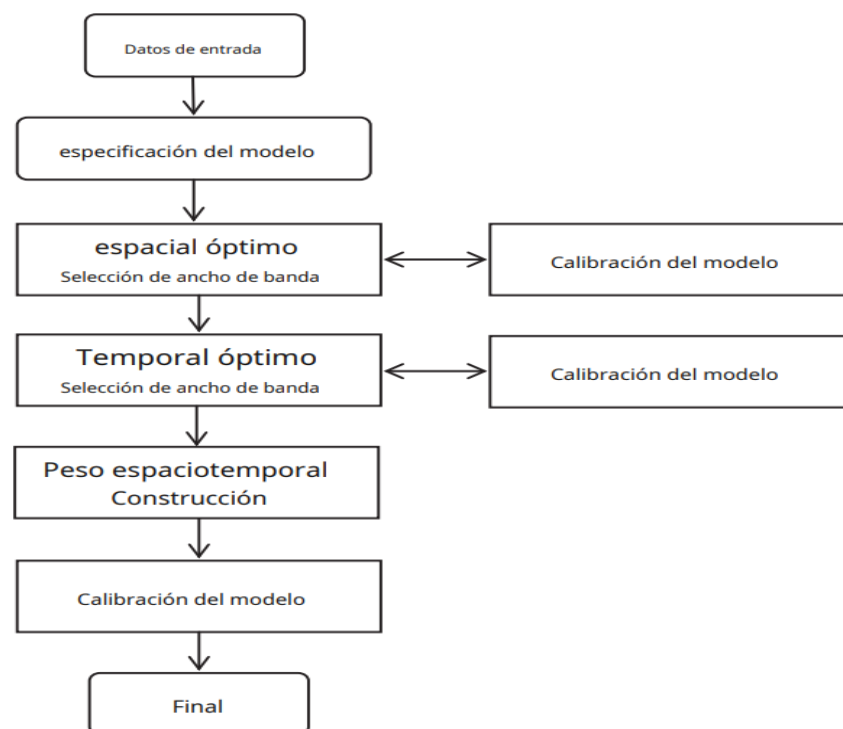


Figura 1.1 Esquema del modelo de regresión geográfico y temporal

La figura 1.1, resume la rutina general para estimar los parámetros locales del GTWR usando la ecuación 1.3. A saber, primero un modelo GWR es especificado usando los datos de entrada en función de la ecuación 1.1, posteriormente se especifica un ancho de banda espacial óptimo para cada período de tiempo, basándose en un criterio de bondad de ajuste

por validación cruzada (CV) o el criterio de información de Akaike (AIC). Luego, utilizando el ancho de banda espacial óptimo se determina el ancho de banda temporal óptimo y éste nuevamente se hace ocupando un criterio de bondad de ajuste, ya sea la validación cruzada (CV) o el criterio de información de Akaike (AIC).

Una vez obtenidos los anchos de banda temporal y espacial óptimos, estos se utilizan para calcular la matriz espaciotemporal de pesos W . Esta matriz es construida a partir de utilización de una función kernel, que por lo general puede ser una función kernel gaussiana o bicuadrada, y son quienes permiten la obtención de las relaciones espacio temporales entre las observaciones; esto para poder estimar los parámetros locales mediante la ecuación 1.3. Es importante resaltar, que tanto para el ancho de banda espacial como para el temporal el cálculo es bastante pesado, pues requiere de repetidas calibraciones temporales del modelo.

Particularmente, para la obtención de la matriz W con los pesos espacio-temporales, una función kernel espacio-temporal es incorporada en el proceso, la cual, se caracteriza por tener anchos de banda espaciales mixtos y un decaimiento a nivel temporal entre las observaciones. Es importante resaltar que en ese tipo de kernel mixto, los pesos dados a los puntos de los datos se calculan no solo en función de la distancia entre el punto de regresión y cada punto de la data, sino también en función de la separación temporal entre ellos. Una expresión que permite obtener dichos pesos uno a uno, se propone desde Fotheringham et al. (2015a) con la siguiente ecuación:

$$w_{ijS,T}^t = k_S(d_{S_{ij}}, b_S) \times k_T(d_{T_{ij}}, b_T). \quad (1.4)$$

La ecuación 1.4 representa la función de ponderación espacio-temporal para cada $w_{ijS,T}^t$, es decir, el peso para cada punto de la data en un instante de tiempo t , con una función kernel espacio-temporal genérica, k_S el kernel espacial y además $d_{S_{ij}}$ es la distancia espacial euclidiana entre el punto de regresión i y un punto j de los datos. Note que cada punto j de los datos, puede estar en cualquier conjunto $S_t, S_{t-1}, \dots, S_{t-q}$. Luego b_S , es el ancho de banda espacial, k_T es el kernel temporal, $d_{T_{ij}}$ es la distancia temporal entre el punto de regresión i y el punto j y b_T es el ancho de banda a nivel temporal. Vale la pena acotar, que con la ecuación 1.4 y desde Fotheringham et al. (2002), suponen que un ancho de banda espacial b_S único se deriva y se aplica a todos los puntos de datos en el conjunto de datos. Sin embargo, no existe razón por la cual se deba suponer previamente un ancho de banda espacial constante en el tiempo. De hecho, también son posibles otras disposiciones de ancho de banda espacial para ajustar los datos. Por ejemplo, parece razonable suponer que los anchos de banda espaciales se vuelvan más pequeños a medida que los puntos de datos se ubican más lejos en el tiempo del punto de regresión. Esto se puede apreciar en la siguiente figura:

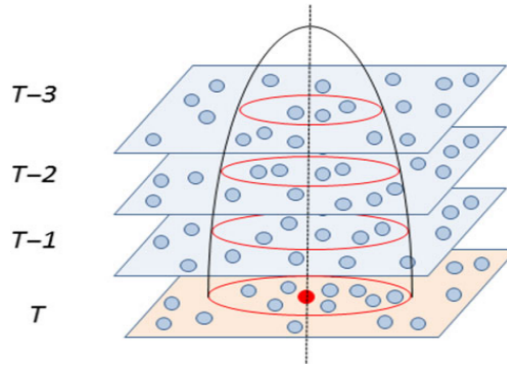


Figura 1.2 Un ejemplo de ancho de banda espaciotemporal de decaimiento en el tiempo.

Note que: 1.2, muestra una descripción para un posible ancho de banda espaciotemporal de decaimiento en el tiempo, donde el período de tiempo de regresión es T y se considera un ancho de banda temporal de 3 unidades, es decir, $(T - 1, T - 2, y T - 3)$. Podemos apreciar, como el ancho de banda espacial se hace más pequeño a medida que las observaciones están más lejos del punto de regresión en el tiempo. Así, el ancho de banda temporal opera de forma similar al espacial, en el sentido de que b_T proporciona cierto tipo de control sobre el rango de amplitud del "círculo de influencia" para los datos geográficos a lo largo del tiempo y que se muestran en color rojo.

Alternativamente se podría pensar lo contrario, es decir, que los anchos de banda espaciales aumentan a medida que los puntos de datos se ubican más lejos en el tiempo con respecto al punto de regresión. Visto de esa manera se estaría compensando el peso temporal más bajo, que el ancho de banda de decaimiento temporal le otorga a las observaciones.

El siguiente paso, es utilizar un conjunto de anchos de banda espaciales segregados a lo largo del tiempo, que deben estimarse junto con el ancho de banda temporal para ajustar los datos. Así, la ecuación 1.4 se extiende a la ecuación 1.5, en donde son posibles diferentes anchos de banda para cada período de tiempo:

$$w_{ijS,T}^t = k_S(d_{Sij}, b_{S_t}) \times k_T(d_{Tij}, b_T) \quad (1.5)$$

En la ecuación 1.5 todos los términos son los mismos que los de la ecuación 1.4, excepto por el término b_{S_t} , que es el ancho de banda específico para un tiempo t . Es decir, si se ocupa la estructura de ponderación que derivada de la ecuación 1.4, la matriz W que se genera, es una matriz diagonal de orden $(n_{\Sigma_T} \times n_{\Sigma_T})$, donde n_{Σ_T} es el número total de tiempos $t, \dots, t - q$ usados para calibrar el modelo. Sin embargo, como se mencionó anteriormente ello supondría un ancho de banda espacial b_s único que se aplicaría a todas las observaciones del conjunto de datos y que no necesariamente es cierto, pues usando el esquema de ponderación de la

o aplicando (AIC) de corrección, con la finalidad de obtener un conjunto de estimaciones locales optimamente compensadas en cuanto al sesgo y la varianza. Por ejemplo, podemos usar la función (CV) para obtener b_S^* y b_T^* :

$$CV(b_S, b_T) = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i(b_S, b_T))^2}{n}}, \quad (1.6)$$

donde \hat{y}_i es el valor ajustado para y_i con exclusión del punto i en la calibración del proceso. La expresión 1.6 se puede usar para distintos anchos de banda espacial en distintos instantes de tiempo t , más aún 1.6 puede ser escrita como:

$$CV(b_{S_t}, b_{S_{(t-1)}}, \dots, b_{S_{(t-q)}}, b_T) = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i(b_{S_t}, b_{S_{(t-1)}}, \dots, b_{S_{(t-q)}}, b_T))^2}{n}}. \quad (1.7)$$

En donde los $q + 1$ anchos de banda espacial y el ancho de banda temporal en 1.7, deben ser estimados simultáneamente para minimizar la función de validación cruzada (CV).

En la práctica, el proceso de selección para un óptimo ancho de banda espacial y un óptimo ancho de banda temporal, parte por la selección de algún kernel espaciotemporal. Por ejemplo, si consideramos un kernel espacio-temporal gaussiano, los pesos se expresan por la siguiente ecuación:

$$w_{ijS,T}^t = \exp\left(-\frac{d_{sij}^2}{b_{S_t}^2}\right) * \exp\left(-\frac{d_{tij}^2}{b_T^2}\right). \quad (1.8)$$

Para periodos de tiempo t , los siguientes pasos muestran como calcular los ancho de banda óptimos.

1. Supongamos que el ancho de banda temporal b_T se establece en una unidad temporal (por ejemplo: día, mes, año).
2. El ancho de banda espacial por periodos de tiempo t , y que se denota por (b_{S_t}) , se calcula utilizando puntos de datos del tiempo t . En este caso d_{tij}^2 es cero y el kernel gaussiano espacio-temporal se transforma en $w_{ijS,T}^t = \exp\left(-\frac{d_{sij}^2}{b_{S_t}^2}\right)$. Seguidamente, se utiliza el esquema de ponderación GWR, ésto para calibrar el modelo de regresión en puntos de datos de tiempo t . Luego, por la minimización de la función (CV), la banda espacial óptima $b_{S_t}^*$ es obtenida para los primeros n_t elementos de la diagonal en la matriz de pesos en la figura 1.3 en base a la ecuación 1.8 que se tomó como ejemplo.

3. Para obtener los segundos elementos de la diagonal de la matriz de pesos, se incorpora al modelo los datos de puntos del tiempo $(t - 1)$, de esta manera el GWR se usa para calibrar el modelo de regresión ocupando los datos del tiempo t y el tiempo $t - 1$, haciendo la regresión de los puntos en el tiempo t . Los puntos de datos en el instante t son ponderados acorde al primer conjunto elemento de la diagonal de la matriz obtenida de la ecuación 1.3, lo que significa es que $b_{S_t}^*$ se mantiene fijo. Los puntos de datos en el instante $(t - 1)$ son ponderados con un kernel espacio-temporal gaussiano, el cual está definido en 1.8 y de donde será estimado el ancho de banda espacial optimo $b_{S(t-1)}^*$. Ahora bien, para un tiempo $(t-1)$ la distancia $d_{ij}^2 = 1$, entonces el kernel espaciotemporal ocupado para tiempos de datos $(t - 1)$ se transforma en $w_{ijS,t}^{t-1} = \exp\left(-\frac{d_{ij}^2}{b_{S(t-1)}^2}\right) * \exp\left(-\frac{1}{1}\right)$. De manera similar al punto dos (2.) el optimo ancho de banda $b_{S(t-1)}^*$ es derivado por minimización de la función (CV) para puntos de datos de periodos de tiempo t , así el segundo conjuntos de n_{t-1} elementos de la diagonal de la matriz de ponderaciones de la figura 1.3 es derivado al ingresar $b_{S(t-1)}^*$ en la ecuación 1.8 con $t = t - 1$.
4. Análogamente, el tercer conjunto de elementos de la diagonal de la matriz de pesos se tiene incorporando puntos de datos de $t - 2$. En este caso, los puntos de datos de t y $t - 1$ son ponderados usando los anchos de banda optimos $b_{S(t-1)}^*$ y $b_{S_t}^*$ como se especifica en el punto (2.) y (3.) Es decir, ambas bandas espaciales se mantienen constantes. Luego, el esquema de ponderación que es ocupado para los puntos de datos de $t - 2$ deriva de la ecuación 1.8 desde donde $b_{S(t-2)}^*$ es estimado. Por otra parte, la distancia para $t - 2$ es $d_{ij}^2 = 2$ y al aplicar el kernel gaussiano espacio-temporal para el periodo de tiempo $t - 2$ se convierte en $w_{ijS,t}^{t-2} = \exp\left(-\frac{d_{ij}^2}{b_{S(t-2)}^2}\right) * \exp\left(-\frac{2}{1}\right)$. Y nuevamente, como en el punto anterior, el ancho de banda espacial optimo $b_{S(t-2)}^*$ es derivado de la minimización de la función (CV) para puntos de datos en el instante de tiempo t , finalmente y una vez que el ancho de banda optimo para $t - 2$ se obtiene, el tercer conjunto de datos de la diagonal de la matriz de peso es derivado de la ecuación 1.8 con $t = t - 2$.
5. El proceso que se describe anteriormente, se repite incorporando puntos de datos uno por uno del tiempo $t - 3, t - 4, \dots, t - q$ para derivar anchos de banda espaciales óptimos para estos períodos de tiempo, junto con los correspondientes conjuntos de elementos diagonales de la matriz de peso.
6. Una vez que la matriz diagonal para puntos de datos de t a $t - q$ se obtiene, se usa el GWR para calibrar el modelo en puntos de tiempo t ocupando el esquema

de ponderación entregado por la matriz de pesos de la figura 1.4. A partir de la calibración GWR, se obtiene un puntaje de la (CV) que se especifica para el ancho de banda temporal que se supone que es 1 unidad temporal en la ecuación 1.1. Esta puntuación de (CV) debe denominarse $CV_{b_T=1}$.

7. El procedimiento descrito desde el punto uno (1.) hasta el punto seis (6.) se repite para los $q - 1$ posibles anchos de banda temporales según el número de retrasos en el modelo, es decir para b_T igual a 2, 3, 4, ..., o q unidades en el pasado. Así, por cada ancho de banda temporal utilizado para calibrar el modelo, se obtiene una puntuación CV, digamos, $CV_{b_T=1}, CV_{b_T=2}, CV_{b_T=3}, \dots$, y $CV_{b_T=q}$.
8. El ancho de banda óptimo para (b_T^*) es aquel para el que la puntuación mínima CV es obtenida. La selección de un ancho de banda temporal óptimo produce el conjunto final de anchos de banda espaciales: $[b_{S_t}^*, b_{S(t-1)}^*, b_{S(t-2)}^*, \dots, b_{S(t-q)}^*]_{b_T^*}$.
9. Finalmente, estimaciones locales para puntos ubicados en el tiempo t se estiman utilizando la ecuación 1.3 en la que los elementos diagonales de W se derivan utilizando el esquema de ponderación dado por el ancho de banda temporal y espacial óptimo del punto ocho (8.)

El GTWR una vez presentado a la comunidad científica, atrajo la atención de muchos investigadores, pues permitió reformular nuevas ideas de implementación, que no eran posibles con modelos similares, incapaces de captar la relación en datos espacio-temporales. Entre algunas de las investigaciones cabe mencionar, por ejemplo, Miftahus Sholihin (2017) en donde se muestra como el GTWR es mejor que el método de regresión global en la modelización del crecimiento económico en Java Central en 2011-2015. Particularmente, muestran la consistencia fuerte del estimador de mínimos cuadrados, pero con la salvedad de que suponen las covariables independientes entre sí. Otro ejemplo se tiene en Harianto et al. (2021), donde el modelado del GTWR en casos de pobreza en 38 regencias/ciudades en la provincia de Java Oriental durante 2010-2015, mostró la igualdad en la estimación de los resultados al comparar una misma locación, pero usando dos funciones kernel distintas. Seguidamente, en el artículo Peng et al. (2019), el uso del GTWR conllevó a la construcción de un nuevo modelo capaz de reducir la escala espacial en datos de temperatura de la superficie de los suelos, específicamente del espectrorradiómetro de imágenes de resolución moderada de 1000 a 100 metros.

Finalmente, mencionar que también existen otras investigaciones en donde se toma como inspiración el GTWR. Tal es el caso de Wu et al. (2014). En ese trabajo se realizan mejoras

sobre la estructura del modelo, con la finalidad de captar elementos de autocorrelación presentes en los datos. Básicamente, a partir del modelo de regresión ponderado geográficamente (GWR) y el modelo de regresión ponderado geográfica y temporalmente (GTWR), se obtiene un modelo autorregresivo ponderado geográfica y temporalmente (GTWAR), el cual toma en cuenta los efectos no estacionarios y autocorrelacionados en forma simultánea. Allí, se formulan dos marcos de mínimos cuadrados por etapa para estimar el nuevo modelo. Cabe destacar, que se lleva a cabo un estudio de caso sobre los precios de la vivienda en la ciudad de Shenzhen, China, del año 2004 al 2008, con la intención de contrastar los resultados obtenidos entre el GTWR, GWR y el modelo GTWAR.

1.2. Modelo Auto-Regresivo Ponderado Temporal y Geográficamente - (GTWAR)

El modelo autoregresivo ponderado geográfica y temporalmente o (**Geographically and Temporally Weighted Autoregressive model**) GTWAR, es una técnica que inicialmente deriva del modelo GWR y del modelo GTWR, pues por una parte hereda la estructura del modelo original GWR, manteniendo el supuesto de independencia entre las covariables, y por otro lado adopta la metodología del GTWR, capturando la heterogeneidad espacial y temporal mediante la matriz diagonal de pesos. De esta manera, para obtener del modelo GTWAR, se hace una mejora en la matriz de ponderaciones, específicamente un cambio en la estructura de la operación ”+” definida desde el GTWR entre las distancias temporal y espacial e incorporando la información autocorrelacionada en el modelo. Esto último, permite captar de mejor manera los efectos que se producen cuando la distancia temporal y espacial varían en conjunto. Más aún, se puede modelar la interacción entre ambas sin asumir un sistema de coordenadas espacio temporal ortogonal. Ver Wu et al. (2014). Es importante resaltar que, los autores en los trabajos de Wu et al. (2014) y previo a definir un modelo autoregresivo ponderado temporal y geográficamente GTWAR, plantean dentro del mismo trabajo, un modelo mejorado del GTWR, al cual llaman IGTWR, **Improved Geographically and Temporally Weighted Regression model**. La manera como obtiene este modelo, se mencionó anteriormente refiriendonos a los cambios sobre la operación ”+” definida entre las distancias espacial y temporal en el modelo GTWR propuesto en Fotheringham et al. (2015a). Luego, una vez obtenida esta mejora con el modelo IGTWR, se pasa a la construcción del GTWAR incorporando la autocorrelación espacial del mismo. Es decir, se considera la autocorrelación espacial con la introducción de un retraso espacial de la forma $\sum_{j=1}^n \bar{W}_{ij} y_j$ en el modelo de regresión lineal propuesto en Can and Megbolugbe (1997), para

generar la siguiente expresión

$$y_i = \rho_i \sum_{j=1}^n \bar{W}_{ij} y_j + \varepsilon_i \quad i = 1, \dots, n \quad (1.9)$$

donde ρ_i representa un parámetro autorregresivo espacial que varía según las ubicaciones geográficas, ε_i denota el error aleatorio en la relación y \bar{W}_{ij} es el elemento en la i_{th} -ésima fila de la j_{th} -ésima columna de la matriz espacial de pesos \bar{W} con $\bar{W}_{ii} = 0$. Cabe destacar, que los elementos de \bar{W}_{ij} son típicamente normalizados por fila tal que por cada i , la $\sum_{i=1}^n \bar{W}_i = 1$. Consecuentemente el retraso espacial puede ser interpretado como un promedio ponderado de los vecinos. Note que el modelo 1.9 se basa en un modelo local autoregresivo. Ahora, incorporando en 1.9 el modelo IGTWR y en consideración del trabajo en Wu et al. (2014), se define el modelo GTWAR mediante la expresión:

$$y_i = \beta_0(u_i, v_i, t_i) + \rho(u_i, v_i, t_i) \sum_{j=1}^n \bar{W}_{ij} y_j + \sum_{k=1}^d \beta_k(u_i, v_i, t_i) x_{ki} + \varepsilon_i \quad i = 1, \dots, n \quad (1.10)$$

donde $\rho(u_i, v_i, t_i)$ es un parámetro escalar espacio-temporal autoregresivo en el punto i . Aquí, es importante resaltar que las covariables consideradas desde el modelo GTWR son independientes entre si con media y varianza finita, lo que permite sin pérdida de generalidad aplicar el Teorema del Límite Central generalizado Cam (1986) sobre la suma de las covariables expresada en dicho modelo, asegurando que la distribución de la suma sobre las x_{ik} , cuando es lo suficientemente grande para i , sigue una distribución normal. Así, es posible asegurar en el análisis de regresión, que la matriz de covarianza asociada a modelo lineal general posee independencia de los errores. Más aún, los ε_i son independientes de las covariables y estos a su vez son independientes entre si, pero siguiendo una misma distribución. Luego para el modelo GTWAR, a pesar de que se suma un factor autoregresivo de retardo entre las observaciones, también se garantiza que los ε_i son independientes e idénticamente distribuidos.

Ahora bien, para seleccionar el modelo se debe resaltar que la estimación de los parámetros tanto en el GTWR como en el GTWAR es altamente dependiente del ajuste sobre los parámetros μ^S , μ^T y/o el v asociado con la función de ponderación usada. La selección del ajuste de los parámetros para el GTWR y el GTWAR puede ser determinada usando un enfoque de (CV) validación cruzada o el criterio de información de Akaike corregido (AIC) recomendado desde Fotheringham et al. (2002). Luego como el kernel adaptativo usa el mismo número q de vecinos más cercanos, hay un parámetro más que se debe calibrar, así para tener un valor óptimo de q , es posible usar el enfoque de (CV) o el (AIC) corregido. Por

otra parte, desde Fotheringham et al. (2015a) se argumenta que sólo el parámetro de radio $\tau = \frac{\mu^T}{\mu^S}$ juega un papel importante en la construcción de los pesos. Así, los investigadores en dicho trabajo configuran $\mu^S = 1$ para reducir el número de parámetros que se deben ajustar a sólo los tres parámetros q , μ^S y v .

La función de Validación Cruzada (CV) se define como:

$$CV(\lambda) = \sum_{i=1}^n (y_i - \hat{y}^{(-i)}(\lambda))^2 \quad (1.11)$$

donde λ es un vector de parámetros ajustables asociados a los modelos GTWAR y GTWR y $\hat{y}^{(-i)}$ es el valor ajustado de $y_i(\lambda)$ con la observación i omitida del proceso de calibración. Ahora, la función (AIC) de corrección desde Hurvich et al. (1998) se define de la manera siguiente:

$$AIC_c(\lambda) = n \log \left(\frac{RSS(\lambda)}{n} \right) + n \log(2\pi) + n \left(\frac{n + tr(\mathbf{H}(\lambda))}{n - 2 - tr(\mathbf{H}(\lambda))} \right) \quad (1.12)$$

donde n es el número de observaciones de la data, RSS es la suma de los residuos al cuadrado definida como $RSS(\lambda) = \sum_{j=1}^n (y_j - \hat{y}^{(-j)})^2$, y $tr(\mathbf{H}(\lambda))$ es la traza de la matriz $\mathbf{H}(\lambda)$ asociada al modelo GTWR o al modelo GTWAR, y λ un parámetro vectorial que se calibra, el cual satisface la relación $\hat{y} = \mathbf{H}(\lambda)y$. Para la obtención de la matriz de pesos relacionada con el GTWR ver Wu et al. (2014). En la expresión 1.12 el log denota el logaritmo natural, además los parámetros ajustables son obtenidos automáticamente mediante la técnica de minimización de la función (CV) 1.11 o mediante la función (AIC) de corrección presentada en 1.12.

Por otra parte, conocidos los detalles del modelo GTWAR, es importante mencionar algunas de las investigaciones en donde se ha implementado este modelo, tal es el caso Fotheringham et al. (2015a) en donde su implementación en los precios de la vivienda en la ciudad de Shenzhen, China, del año 2004 al 2008, pues los autores demuestran que existen beneficios sustanciales al modelar los efectos de la no estacionariedad espaciotemporal y de la autocorrelación simultáneamente en los precios de la vivienda en términos de R^2 y el criterio de información de Akaike (AIC). Se evidencia una reducción en los errores absolutos en un 31,8% y un 67,7% en relación con los modelos GTWR y GWR, respectivamente, en el conjunto de datos de Shenzhen. Adicionalmente, entre otras de las investigación a mencionar, en las que se vincula el modelo GTWAR está Djalihu et al. (2020). Aquí, los autores considerando que en el modelo GTWR no todas las variables eran locales o insignificantes (de naturaleza global), se desarrolla un modelo MGTWR (**Mixed Geographically and Temporally Weighted Regression**), es decir, una regresión mixta ponderada geográfica

y temporalmente que junto al GTWAR hizo posible la adaptación de un modelo mixto autoregresivo ponderado geográfica y temporalmente. Dicho modelo MGTWAR (**Mixed Geographically and Temporally Weighted Autoregressive**), permitió modelar datos sobre el porcentaje de personas pobres en la regencia/municipio de la isla de Java en 2012-2018, evidenciando que el modelo MGTWAR produjo un criterio de información de Akaike (AIC) más pequeño que GTWR. Otras investigación, además de las mencionadas, en dónde se implementa y vincula al modelo

GTWAR, se pueden ver en Shim and Hwang (2018) y Luo et al. (2021). Gracias a su capacidad predictiva a nivel espacial y temporal en datos georeferenciados, lo convirtió en una herramienta muy útil para los investigadores en estos trabajos.

1.3. Ecuaciones en derivadas parciales estocásticas (SPEDs)-dirigidas por un ruido coloreado fraccionario

En la presente sección, revisaremos algunos trabajos principalmente asociados a ecuaciones en derivadas parciales estocásticas impulsadas por un ruido fraccionario coloreado, es decir, un ruido que es fraccionario en tiempo con índice de Hurst $H \in (\frac{1}{2}, 1)$ y coloreado en espacio con una estructura de covarianza espacial del tipo Gaussiano. Iniciaremos por entregar las condiciones de existencia y unicidad de soluciones en las SPEDs que son perturbadas por un ruido, en particular la ecuación del calor. Para esto, nos apoyaremos en los trabajos de Balan and Tudor (2007) posteriormente conectaremos con una parte de los resultados expuestos en Torres et al. (2014) y finalmente describir las técnicas y resultados expuestos en Gyöngy (1999), las cuales permiten discretizar las SPDEs.

Considerando la siguiente ecuación introducida desde Balan and Tudor (2007):

$$u_t - \Delta u = \dot{B}, \quad u \in ((0, T) \times \mathbb{R}^d),$$

la cual, posee condiciones iniciales nulas $u(0, \cdot) = 0$, donde \dot{B} es un ruido gaussiano fraccionario en el tiempo, (con índice de Hurst $H \in (1/2, 1)$), y coloreado en espacio, con covarianza espacial representada por una función f . Además, u es la solución general, u_t la derivada parcial con respecto al tiempo en $(0, T) \subseteq \mathbb{R}^+$ y Δu la derivada parcial de segundo orden respecto del espacio en \mathbb{R}^d . Mostraremos condiciones sobre la existencia y unicidad de las soluciones de la misma en aras de conectar con otras investigaciones relevantes.

Básicamente, encontraron la condición necesaria y suficiente sobre H para la existencia del proceso solución y una distribución solución de esta ecuación en particular. Lo cual, fue posible dado que desde Dalang (1999) ya se había realizado la construcción del cálculo

que permitía encontrar la solución de SPDEs en donde la función de Green, (Solución fundamental de la ecuación) no era una función propiamente, sino una distribución de Schwartz. Más aún, desde Balan and Tudor (2007) se prueba que para la ecuación del calor, cuando se colorea el espacio con la función f como Kernel de Riesz dado por la ecuación:

$$f_{\alpha}(u) := \gamma_{\alpha,d} \|u\|^{-d+\alpha}, \quad \text{para } -d < \alpha < d \quad \text{con } \gamma_{\alpha,d} = 2^{d-\alpha} \pi^{d/2} \Gamma((d-\alpha)/2) / \Gamma(\alpha/2). \quad (1.13)$$

Como estructura espacial, se puede compensar los inconvenientes que genera el ruido coloreado fraccionario al momento de encontrar la solución para la ecuación del calor. Cabe destacar, que la función kernel 1.13 fue considerada dentro del resultado principal en nuestro artículo Fermín et al. (2023). Esto, para definir la medida del ruido fraccionario coloreado espacio temporal que perturba al modelo que proponemos.

Por otro lado, a partir de los resultados obtenidos por los autores en Torres et al. (2014), se logró estudiar la variaciones cuadráticas de un tipo de procesos Gaussianos que muestra la ecuación del calor. Particularmente, se construye un estimador H que permite estudiar el comportamiento asintótico de las trayectorias exhibidas por procesos autosimilares en el tiempo del tipo mencionado, y que se presentan como soluciones de SPDEs dirigidas por un ruido fraccionario coloreado. Cabe destacar, que esto último garantiza bajo cierto orden de regularidad de las trayectorias, la cercanía al movimiento Browniano fraccionario.

Finalmente, pero no menos importante, mencionar los trabajos realizados en Gyöngy (1999), ya que presentan una manera de discretizar las soluciones de SPDEs. Específicamente, hacen aproximaciones de SPDEs parabólicas cuasi-lineales sustituyendo las derivadas por diferencias finitas y obteniendo las condiciones de convergencia tanto para esquemas implícitos, como explícitos. Cabe mencionar, que para la ecuación del calor estocástica, mediante un esquema de Euler en diferencias finitas y a partir del resultado mencionado, nosotros en el presente trabajo logramos presentar la discretización de las soluciones de la ecuación del calor en estocástica en una forma discretizada, que es en su versión discreta un proceso autoregresivo, del tipo GTWAR.

1.4. Contribuciones

En la presente investigación, se obtienen dos resultados importantes, el primero de ellos es la estimación de un modelo de regresión espacio-temporal ponderado dirigido por un ruido fraccionario coloreado. Es decir, considerando un modelo de regresión GTWR introducido por los trabajos de Fotheringham et al. (2015a), con ecuación 1.1, y dirigido por

un ruido coloreado fraccionario con características y condiciones como las presentada por los autores en Torres et al. (2014). Planteamos un modelo de regresión espacio-temporal ponderado con un ruido coloreado en espacio y fraccionario en tiempo. Ver el siguiente artículo Fermín et al. (2023). El modelo se obtuvo una vez demostrado el principal resultado en la investigación mencionada. Es decir, el teorema 2.1, que permite probar la consistencia sobre el estimador de mínimos cuadrados ponderados $\hat{\beta}(z_i)$, del vector de parámetros locales $\beta(z_i)$, con z_i una observación espacio-temporal. Para esto, nos apoyamos en dos resultados adicionales obtenidos dentro de la misma prueba, es decir, el lema 1 que entrega la definición para la covarianza del ruido espacio-temporal y la proposición 2.3 que contribuye a la prueba de convergencia del denominador de i) dentro del teorema 2.1. Así, haciendo uso de estos resultados que mencionamos, comprobamos la convergencia en probabilidad del estimador $\hat{\beta}(z_i)$ para $2H + d + \alpha > 0$ y $d + 1 + \beta > 0$. En particular, para $H > 0$, $\alpha > -d$, y $\beta > -(d + 1)$. Y la convergencia casi segura, para $2H + d + \alpha - 1 > 0$ y $\beta > 0$, en particular para $H > 1/2$, $\alpha \leq 0$ y $\beta > 0$. Luego, para complementar dichos resultados teóricos, realizamos algunas simulaciones tanto para el modelo obtenido, combinando tanto el H para la covarianza espacial como para la covarianza temporal. Esto varía la rugosidad de la superficie simulada como la larga dependencia del fraccionario en tiempo respectivamente. Además las pruebas de consistencia de las simulaciones del modelo muestran la adecuación del mismo bajo los supuestos considerados.

Ahora bien, como un segundo aporte realizamos una discretización sobre una base finita de autofunciones del operador de segundo orden para la ecuación del calor, apoyándonos en los resultados obtenidos en Gyöngy (1999) y donde planteamos la forma explícita de las soluciones de la ecuación del calor en forma discretizada, las cuales son un modelo autoregresivo ponderado y que se identifican con el modelo autorregresivo geográfico y temporalmente ponderado GTWAR.

La organización de la presente investigación se hizo de acuerdo al orden de los elementos trabajados, es decir, en un capítulo introductorio dejamos una revisión de los elementos teóricos necesarios para la comprensión de los trabajos que se proponen posteriormente. En un segundo capítulo presentamos los resultados del artículo Fermín et al. (2023) iniciando por la introducción del modelo en la primera sección, posteriormente se define el movimiento Browniano fraccionario en una segunda sección, para continuar con la definición del ruido coloreado fraccionario, posterior a ello presentamos el tipo de correlación que suponemos entre las covariables, luego precede el planteamiento del estimador y finalmente la prueba de la consistencia para dicho estimador. Una última sección dentro del segundo capítulo se encuentran las simulaciones. Finalmente, en el tercer capítulo se dejan los resultados obtenidos mediante una discretización de Euler, sobre una base finita de autofunciones para

el operador de segundo sobre la ecuación del calor, lo cual permitió la obtención de las soluciones de la ecuación discretizadas para esto, en donde posteriormente dejamos explícita la forma de las mismas, evidenciando la similitud del modelo autoregresivo obtenido con el modelo de regresión geográfico y temporalmente ponderado.

Capítulo 2

Modelo de regresión geográfica y temporalmente ponderada con ruido coloreado fraccionario

En este capítulo Consideramos un modelo de regresión GTWR introducido por los trabajos de Fotheringham et al. (2015a), con ecuación 1.1, y dirigido por un ruido coloreado fraccionario con características y condiciones como las presentada por los autores en Torres et al. (2014). Planteamos un modelo de regresión geográfica y temporalmente ponderado con un ruido coloreado en espacio y fraccionario en tiempo. Ver el siguiente artículo Fermín et al. (2023). El modelo se obtuvo una vez demostrado nuestro principal resultado. Es decir, el teorema 2.1, que permite probar la consistencia sobre el estimador de mínimos cuadrados ponderados $\hat{\beta}(z_i)$, del vector de parámetros locales $\beta(z_i)$, con z_i una observación espacio-temporal. Para esto, nos apoyamos en dos resultados adicionales. El lema 1 que entrega la definición para la covarianza del ruido espacio temporal y la proposición 2.3 que contribuye a la prueba de convergencia del denominador de i) dentro del teorema 2.1. Así, haciendo uso de estos resultados que mencionamos, comprobamos la convergencia en probabilidad del estimador $\hat{\beta}(z_i)$ para $2H + d + \alpha > 0$ y $d + 1 + \beta > 0$. En particular, para $H > 0$, $\alpha > -d$, y $\beta > -(d + 1)$ y la convergencia casi segura, para $2H + d + \alpha - 1 > 0$ y $\beta > 0$. En particular para $H > 1/2$, $\alpha \leq 0$ y $\beta > 0$.

En el presente trabajo estimamos un modelo de regresión con ruido fraccionario coloreado. Para esto, nos basamos en un modelo espacio-temporal con ruido fraccionario coloreado. Posteriormente se por un ruido fraccionario coloreado. el resultado principal. En este sentido, en la sección 2.1 presentamos el modelo GWRT con ruido coloreado fraccionario explicando cada uno de los elementos que lo componen, asumiendo que la esperanza del ruido dada

las covariables es nula. En la sección 2.2 definimos el movimiento Browniano fraccionario indicando sus propiedades y características que contribuyen al trabajo. Para la sección 2.3 definimos la medida del ruido espacio-temporal apoyándonos en la medida temperada del kernel de Riesz de orden α , la cual asumimos acotada, tal como se plantea los trabajos en Torres et al. (2014) y que nos permitió construir el lema 1 referido a la estructura de covarianza del ruido coloreado fraccionario. Seguidamente para la sección 2.4 indicamos los supuestos sobre el tipo de interacción local, entre las observaciones y entre las covariables. Asumimos que tanto la función de covarianza local de las observaciones como de las covariables son simétricas, definidas positiva y cierto tipo de Hölder-Continuidad. Cabe destacar, que esto nos permitió considerar dos tipos de interacciones, una interacción débil entre las covariables y una fuerte. siendo esta última la escogida para la prueba de consistencia. En la sección 2.5 mostramos el estimador de mínimos cuadrados para la función del modelo 2.1, indicando las condiciones que debe cumplir la matriz 2.26 y las consideraciones para la función kernel en dicha matriz, la cual actúa sobre las observaciones. Finalmente, en la sección 2.6 probamos la proposición 2.3 que contribuye a la prueba del Teorema 2.1, el cual muestra la consistencia del estimador propuesto. Cabe destacar que usamos análisis de variaciones cuadráticas como técnica que deriva de la propuesta en Torres et al. (2014) y que son se desarrollaron a partir de los resultado en Dalang (1999) y Balan and Tudor (2007).

2.1. Modelo GTWR

Consideramos un modelo de regresión ponderado geográfica y temporalmente (GTWR) conducido por un ruido fraccionario coloreado, el cual viene representado por la siguiente ecuación:

$$Y_i = \beta_0(z_i) + \sum_{j=1}^p \beta_j(z_i)X_{i,j} + \varepsilon_i, \quad i = 1, \dots, n; \quad p \in \mathbb{N}, \quad (2.1)$$

donde $z_i = (t_i, u_i)$ denota las coordenadas de la observación en el punto z_i , en el espacio $u_i \in \mathbb{R}^d$ en un tiempo $t_i \in \mathbb{R}^+$, $\beta_0(z_i)$ indica el valor del intercepto, $\beta_j(z_i)$ indica el parámetro asociado a la j th covariable X_j en el punto z_i , y ε_i es el ruido fraccionario coloreado en el punto z_i ; Es decir, $\varepsilon = (\varepsilon_i)_{i=1, \dots, n}$ es un ruido gaussiano el cual se comporta como un Movimiento Browniano fraccionario (mBf) en tiempo y tiene covarianza espacial blanca o coloreada.

Hipótesis 2.1. *Asumimos que $\varepsilon = (\varepsilon_i)_{i=1, \dots, n}$ es independiente de las covariables (X_1, \dots, X_p) .*

2.2. Movimiento Browniano Fraccionario (mBf)

A.N Kolmogorov en Kolmogorov (1940), fue uno de los primeros matemáticos en considerar en 1940 un proceso Gaussiano con incrementos estacionarios y que posee la propiedad de autosimilaridad. Esto último puede ser especificado por:

decimos que un **proceso X es autosimilar con constante de autosimilaridad H** si:

$$\forall a, \exists H > 0 \text{ tal que } \{X_{at}; t \geq 0\} \cong \{a^H X_t; t \geq 0\}, \quad (2.2)$$

donde \cong denota igualdad en distribución.

De esta forma el movimiento Browniano fraccionario B^H es un proceso Gaussiano con media cero, y con función de covarianza de la siguiente forma:

$$E(B_t^H, B_s^H) = \frac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H}) \quad 0 < H < 1 \quad (2.3)$$

El mBf es el único proceso continuo Gaussiano con media cero con incrementos estacionarios y coeficiente de autosimilaridad a^H .

El cálculo estocástico del mBf proviene de un trabajo de Mandelbrot y Van Ness Mandelbrot and van Ness (1968) el cual, considera la representación integral por medias móviles del mBf B^H vía el proceso de Wiener en un intervalo. Ver Biagini et al. (2008), Mishura et al. (2008) y Nourdin (2012).

Definición 2.1. *El Movimiento Browniano Fraccionario (mBf) con índice $H \in (0, 1)$ es un proceso Gaussiano $B^H = (B_t^H; t \in \mathbb{R}^+)$ sobre el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ con las siguientes propiedades:*

i) $B_0^H = 0$

ii) $E(B_t^H) = 0 \quad t \in \mathbb{R}^+$

iii) $E(B_t^H B_s^H) = \frac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H}), \quad t, s \in \mathbb{R}^+$

Cabe destacar, que cuando $H = \frac{1}{2}$ el mBf es el movimiento Browniano estándar, pero para $H \neq \frac{1}{2}$ el proceso no es semimartingala ni tiene la propiedad de Markov.

Observación 2.1. *De la definición anterior se deduce que: Cuando un mBf es autosimilar con parámetro de autosimilaridad H , significa que $\{B_{ct}^H\} \cong \{c^H B_t^H\}$, para todo $c > 0$.*

Observación 2.2. *El mBf tiene incrementos estacionarios en el siguiente sentido, $E(B_t^H - B_s^H)^2 = |t - s|^{2H}$, es decir, depende solamente de la diferencia de los tiempos. Más aún las trayectorias del mBf son Hölder continuas.*

Definición 2.2. Una función $f : [a, b] \rightarrow \mathbb{R}$ es Hölder continua con exponente $\lambda \in (0, 1]$, si existe una constante $c > 0$ tal que $\forall s, t \in [a, b]$, $|f(t) - f(s)| \leq c|t - s|^\lambda$.

Ahora bien, cabe mencionar que la demostración de la Hölder continuidad del mBf está basada en el siguiente teorema de continuidad de Kolmogorov.

Proposición 2.1 (Kolmogorov). Sea $\{X_t : t \in [0, T]\}$ un proceso estocástico separable satisfaciendo que existen constantes $k > 0$, $\alpha > 0$ $\beta > 0$ tal que

$$E|X_t - X_s|^\alpha \leq k|t - s|^{1+\beta}, \quad 0 \leq s < t \leq T, \quad (2.4)$$

entonces X es un proceso continuo.

Más aún, si el proceso estocástico $\{X_t / t \in [0, T]\}$ es separable, y satisface la condición 2.4 entonces las trayectorias son Hölder continuas hasta el orden β/α .

Observación 2.3. De lo anterior, tomando en cuenta que B^H es Gaussiano y ocupando el teorema (Kolmogorov), podemos deducir que:

i) Existe una modificación para trayectorias continuas.

ii) Casi todas las trayectorias de B^H son localmente Hölder continuas hasta el orden H .

Proposición 2.2. El proceso Gaussiano $\{X_t / t \in [0, T]\}$ es Hölder continuo hasta orden H . Es decir, $\exists c_\varepsilon$ variable aleatoria tal que

$$|X_t - X_s| \leq c_\varepsilon |t - s|^{H-\varepsilon}, \quad \forall \varepsilon \in (0, H), \quad (2.5)$$

si y sólo si existe una constante c_ε tal que $E(X_t - X_s)^2 \leq c_\varepsilon |t - s|^{2H-\varepsilon}$, $\forall \varepsilon \in (0, H)$. Ver Azmoodeh et al. (2014)

Observación 2.4. La variable aleatoria c_ε en 2.5 satisface que $E(\exp\{\alpha c_\varepsilon^k\}) < \infty$, para toda constante $\alpha \in \mathbb{R}$ y $k < 2$; y también para $k = 2$ con α pequeño. En particular, los momentos de todo orden de c_ε son finitos.

Consecuentemente $\forall 0 < \varepsilon < H$ y $T > 0$ existe una variable aleatoria positiva $G_{\varepsilon, T}$ tal que $\forall s, t \in [0, T]$, ocurre que el

$$\sup_{s, t \in [0, T]} |B_t^H - B_s^H| \leq G_{\varepsilon, T} |t - s|^{H-\varepsilon} \text{ a.s.} \quad (2.6)$$

Más aún, de acuerdo a la prueba del Lema (7.4) en Nualart David Rascanu (2002), es posible elegir $G_{\varepsilon, T} = c_{\varepsilon, T} \varepsilon_1$ donde $c_{\varepsilon, T}$ es una constante y ε_1

$$\varepsilon_1 = \left(\int_0^T \int_0^T \frac{|B_r^H - B_u^H|^{\frac{\varepsilon}{2}} dr du}{|r-u|^{\frac{2H}{\varepsilon}}} \right)^{\frac{\varepsilon}{2}}, \quad (2.7)$$

tal que $E(|G_{\varepsilon,T}|^P) < \infty$, para todo $P \geq 1$.

Ahora bien, los incrementos del movimiento Browniano fraccionario (mBf), tal cómo se definió en 2.1, son independientes en el caso $H = \frac{1}{2}$. Son negativamente correlacionados para $H \in (0, 1/2)$ y positivamente correlacionados para $H \in (1/2, 1)$. Más aún, un mBf con $H \in (1/2, 1)$ posee la propiedad de memoria larga, lo cual significa que:

$$\sum_{n=1}^{\infty} r(n) = \infty \quad (2.8)$$

donde

$$r(n) = E(B_1^H (B_{(n+1)}^H - B_n^H)) > 0 \quad (2.9)$$

es su función de autocovarianza. Ver Mandelbrot and van Ness (1968).

2.3. Ruido Coloreado Fraccionario

Iniciaremos con una breve descripción de la covarianza espacial del ruido y para ello, recordemos los trabajos de Torres et al. (2014). Consideremos μ una medida temperada no negativa sobre \mathbb{R}^d , es decir, una medida que satisface:

$$\int_{\mathbb{R}^d} (1 + |\xi|^2)^{-\ell} \mu(d\xi) < \infty, \quad \ell > 0 \quad (2.10)$$

Observación 2.5. *Asumiremos que 2.10 se cumple en el desarrollo de lo que sigue más adelante.*

Hipótesis 2.2.

$$\int_{\mathbb{R}^d} (1 + |\xi|^2)^{-\ell} \mu(d\xi) < \infty, \quad \ell > 0$$

Observación 2.6. *Note que $\xi \in \mathbb{R}^n$, entonces podemos escribir $\xi = (\xi_1, \xi_2, \dots, \xi_n)$. Además, $|\xi| = |(\xi_1, \xi_2, \dots, \xi_n)|$ es la norma Euclídea sobre \mathbb{R}^n aplicada a ξ , por lo tanto $(1 + |\xi|^2)$ se comporta como una constante alrededor del cero. Esto significa, que si se toman elementos de \mathbb{R}^n cercanos al origen o al cero de este espacio y se sustituyen en la expresión $(1 + |\xi|^2)$, se obtiene como resultado valores de la norma cada vez más pequeños en \mathbb{R} , los cuales al*

ser elevados al cuadro resultan en valores aún más cercanos al cero de los números reales y que finalmente al sumar con 1 es casi 1 nuevamente, es decir, un valor constante. Por otra parte, si ocurre lo contrario y se toman valores más alejados del cero en el espacio \mathbb{R}^n , entonces los valores que se obtienen de aplicar la norma son números grandes y por ende aún más grandes cuando se elevan al cuadrado, por lo que en la expresión $(1 + |\xi|^2)$ el aporte de 1 en la suma es casi despreciable, predominando el valor obtenido de $|\xi|^2$, es decir, se comporta como $|\xi|^2$ a medida que los valores son más alejados del cero o cada vez que tienden al infinito. Luego, la condición anterior sobre μ es equivalente a $\int_{|\xi| \leq 1} \mu(d\xi) < \infty$ y $\int_{|\xi| \geq 1} \frac{1}{\|\xi\|^{2\ell}} \mu(d\xi) < \infty$, para algún $\ell \geq 1$.

Consideremos a $f : \mathbb{R}^d \rightarrow \mathbb{R}^+$ la transformada de Fourier de μ en $S'(\mathbb{R}^d)$, es decir:

$$\int_{\mathbb{R}^d} f(u) \varphi(u) du = \int_{\mathbb{R}^d} \mathcal{F} \varphi(\xi) \mu(d\xi), \quad \forall \varphi \in S(\mathbb{R}^d), \quad (2.11)$$

donde $\mathcal{F} \varphi$ es la transformada de Fourier de la función φ , $S(\mathbb{R}^d)$ es el espacio de funciones de Schwartz de todas las funciones decrecientes en \mathbb{R}^n y $S'(\mathbb{R}^d)$ el espacio de las distribuciones temperadas, es decir, funcionales lineales continuos en $S(\mathbb{R}^d)$.

Sea H el parámetro de Hurst en $(1/2, 1)$. Sobre un espacio de probabilidad completo $(\Omega, \mathcal{F}, \mathbb{P})$, consideremos un campo gaussiano de media cero $W^H = \{W_t^H(A) : t \geq 0, A \in \mathcal{B}_b(\mathbb{R}^d)\}$, definido sobre el conjunto de las funciones acotadas Borel- medibles $\mathcal{B}_b(\mathbb{R}^d)$, con función de covarianza:

$$\mathbb{E} (W_t^H(A) W_s^H(B)) = R_H(t, s) \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{I}_A(u) f(u-v) \mathbb{I}_B(v) dudv := \langle \mathbb{I}_{[0,t] \times A}, \mathbb{I}_{[0,s] \times B} \rangle_{\mathcal{H}}, \quad (2.12)$$

donde R_H es la covarianza del mBf.

$$R_H(t, s) = \frac{1}{2} (t^{2H} + s^{2H} - |t-s|^{2H}), \quad \text{for } s, t \geq 0. \quad (2.13)$$

Esto, se puede extender a una medida del ruido gaussiano en $\mathcal{B}_b(\mathbb{R}^+ \times \mathbb{R}^d)$ definiendo los incrementos del mBf en espacio-tiempo como se plantean a continuación:

$$W^H((s, t] \times A) := W_t^H(A) - W_s^H(A). \quad (2.14)$$

Cabe destacar que los detalles de cómo se realiza la medida sobre 2.14 se muestran en la siguiente sección, al igual que la obtención de la función de covarianza para el ruido coloreado o ruido espacio-temporal mediante el lema 1, que también es demostrado en la presente sección.

Supongamos que la covarianza del ruido espacial es la función kernel de Riesz de orden α , la cual satiface el siguiente supuesto:

Hipótesis 2.3. $f_\alpha(u) := \gamma_{\alpha,d} \|u\|^{-d+\alpha}$, for $-d < \alpha < d$ and $\gamma_{\alpha,d} = 2^{d-\alpha} \pi^{d/2} \Gamma((d-\alpha)/2) / \Gamma(\alpha/2)$. En este caso, $\mu(d\xi) = \|\xi\|^{-\alpha} d\xi$.

Observación 2.7. Considerando el supuesto 2.3, el supuesto 2.2 se satisface para $d - \alpha < 2\ell$. Cabe destacar, que el caso especial de ruido blanco en espacio es idénticamente igual al caso particular del supuesto 2.3 con $\alpha = 0$, en donde μ es la medida Lebesgue.

El ruido fraccionario coloreado en un punto de observación z_l se define como $\varepsilon_l = \Delta W^H(z_l)$, lo cual representa la medida del ruido en un entorno $V(z_l)$, definido por

$$V(z_l) = \left\{ z \in \mathbb{R}^+ \times \mathbb{R}^d : \|z - z_l\| = \max\{|t - t_l|, \|u - u_l\|\} \leq \delta_n \right\},$$

donde δ_n es tal que el volumen del entorno $V(z_l)$ es $1/n$; Es decir,

$$\lambda(V(z_l)) = \int_{\mathbb{R}^+} \int_{\mathbb{R}^d} \mathbb{I}_{\{|t-t_l| \leq \delta_n\}} \mathbb{I}_{\{\|u-u_l\| \leq \delta_n\}} dt du = 2\delta_n (\delta_n)^d \lambda(\mathcal{S}^{d-1}) = \frac{1}{n}, \quad (2.15)$$

con $\lambda(\mathcal{S}^{d-1})$ el volumen de una hiperesfera de radio unitario de dimensión d .

Ahora, si reescribimos $V(z_l)$ como $V(z_l) = (t_l^-, t_l^+] \times V(u_l)$, donde $t_l^\pm = t_l \pm \delta_n$ y $V(u_l) = \{u \in \mathbb{R}^d : \|u - u_l\| \leq \delta_n\}$. Entonces,

$$\varepsilon_l = \Delta W^H(z_l) := W^H(V(z_l)) = W_{t_l^+}^H(V(u_l)) - W_{t_l^-}^H(V(u_l)). \quad (2.16)$$

Lo que representa los incrementos del ruido fraccionario coloreado a partir de la medida del ruido en un entorno de z_l . Más aún, se puede representar la covarianza del ruido fraccionario coloreado

$\varepsilon = (\varepsilon_l)_{l=1:n}$ mediante la función del lema 1 que se deja a continuación.

Lemma 1. La función de covarianza del ruido coloreado fraccionario $\varepsilon = (\varepsilon_l)_{l=1:n}$ está definida por

$$\mathbb{E}(\varepsilon_l \varepsilon_{l'}) = \frac{1}{2} \left(\int_{t_l^-}^{t_l^+} \int_{t_{l'}^-}^{t_{l'}^+} 2H(2H-1) |t-t'|^{2H-2} dt' dt \right) \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{I}_{V(u_l)}(u) \gamma_{\alpha,d} \|u-v\|^{-d+\alpha} \mathbb{I}_{V(u_{l'})}(v) dudv \right) \quad (2.17)$$

y su varianza es $\mathbb{E}(\varepsilon_l^2) = \sigma^2 2^{2H} (\delta_n)^{2H+d+\alpha}$, con $\sigma^2 = \text{Var}(W^H(\mathbb{I}_{\{0 \leq t \leq 1, \|u\| \leq 1\}}))$.

Demostración. De las ecuaciones 2.12 y 2.16 se tiene que la función de covarianza de $\varepsilon = (\varepsilon_l)_{l=1:n}$ es

$$\begin{aligned}
\mathbb{E}(\varepsilon_l \varepsilon_{l'}) &= \mathbb{E} \left(W^H(V(z_l)) W^H(V(z_{l'})) \right) \\
&= \mathbb{E} \left(\left(W_{t_l^+}^H(V(u_l)) - W_{t_l^-}^H(V(u_l)) \right) \left(W_{t_{l'}^+}^H(V(u_{l'})) - W_{t_{l'}^-}^H(V(u_{l'})) \right) \right) \\
&= \frac{1}{2} \left[|t_l - t_{l'} + 2\delta_n|^{2H} + |t_l - t_{l'} - 2\delta_n|^{2H} - 2|t_l - t_{l'}|^{2H} \right] \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{I}_{V(u_l)}(u) f(u-v) \mathbb{I}_{V(u_{l'})}(v) dudv \\
&= \frac{1}{2} \left(\int_{t_l^-}^{t_l^+} \int_{t_{l'}^-}^{t_{l'}^+} 2H(2H-1) |t-t'|^{2H-2} dt' dt \right) \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{I}_{V(u_l)}(u) \gamma_{\alpha,d} \|u-v\|^{-d+\alpha} \mathbb{I}_{V(u_{l'})}(v) dudv \right).
\end{aligned} \tag{2.18}$$

de donde en 2.11 se tiene que la función de covarianza espacial puede ser reescrita como

$$\begin{aligned}
&\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{I}_{V(u_l)}(u) f(u-v) \mathbb{I}_{V(u_{l'})}(v) dudv \\
&= \int_{\mathbb{R}^d} \mathcal{F} \mathbb{I}_{V(u_l)}(\xi) \overline{\mathcal{F} \mathbb{I}_{V(u_{l'})}(\xi)} \mu(d\xi) \\
&= \int_{\mathbb{R}^d} \left(\int_{\|u-u_{l'}\| \leq \delta_n} e^{-i\xi \cdot u} du \right) \left(\int_{\|v-u_{l'}\| \leq \delta_n} e^{i\xi \cdot v} dv \right) \mu(d\xi) \\
&= \int_{\mathbb{R}^d} \left((\delta_n)^d \int_{\|u\| \leq 1} e^{-i\xi \cdot (u_l + \delta_n u)} du \right) \left((\delta_n)^d \int_{\|v\| \leq 1} e^{i\xi \cdot (u_{l'} + \delta_n v)} dv \right) \mu(d\xi) \\
&= (\delta_n)^{2d} \int_{\mathbb{R}^d} e^{-i\xi \cdot (u_l - u_{l'})} \left(\int_{\|u\| \leq 1} e^{-i\xi \cdot \delta_n u} du \right) \left(\int_{\|v\| \leq 1} e^{i\xi \cdot \delta_n v} dv \right) \mu(d\xi) \tag{2.19} \\
&= (\delta_n)^{2d} \int_{\mathbb{R}^d} e^{-i\xi \cdot (u_l - u_{l'}) / \delta_n} \left(\int_{\|u\| \leq 1} e^{-i\xi u} du \right) \left(\int_{\|v\| \leq 1} e^{i\xi v} dv \right) \left\| \frac{\xi}{\delta_n} \right\|^{-\alpha} (\delta_n)^{-d} d\xi \\
&= (\delta_n)^{d+\alpha} \int_{\mathbb{R}^d} e^{-i\xi \cdot (u_l - u_{l'}) / \delta_n} \left(\int_{\|u\| \leq 1} e^{-i\xi u} du \right) \left(\int_{\|v\| \leq 1} e^{i\xi v} dv \right) \|\xi\|^{-\alpha} d\xi \\
&= (\delta_n)^{d+\alpha} \int_{\mathbb{R}^d} \left(\int_{\|u-u_{l'}/\delta_n\| \leq 1} e^{-i\xi u} du \right) \left(\int_{\|v-u_{l'}/\delta_n\| \leq 1} e^{i\xi v} dv \right) \mu(d\xi) \\
&= (\delta_n)^{d+\alpha} \text{Cov} \left(\mathbb{I}_{\{\|u-u_{l'}/\delta_n\| \leq 1\}}, \mathbb{I}_{\{\|u-u_{l'}/\delta_n\| \leq 1\}} \right),
\end{aligned}$$

en particular, para $l = l'$ obtenemos la varianza espacial

$$\begin{aligned}
& \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{I}_{V(u_l)}(u) f(u-v) \mathbb{I}_{V(u_l)}(v) du dv \\
&= (\delta_n)^{d+\alpha} \int_{\mathbb{R}^d} \left(\int_{\|u\| \leq 1} e^{-i\xi u} du \right) \left(\int_{\|v\| \leq 1} e^{i\xi v} dv \right) \|\xi\|^{-\alpha} d\xi \\
&= (\delta_n)^{d+\alpha} \int_{\mathbb{R}^d} \left(\int_{\|u\| \leq 1} e^{-i\xi u} du \right) \left(\int_{\|v\| \leq 1} e^{i\xi v} dv \right) \mu(d\xi) \\
&= (\delta_n)^{d+\alpha} \text{Var}(\mathbb{I}_{\{\|u\| \leq 1\}}) \\
&= \sigma^2 (\delta_n)^{d+\alpha},
\end{aligned} \tag{2.20}$$

donde $\sigma^2 = \text{Var}(\mathbb{I}_{\{\|u\| \leq 1\}})$. Así, la varianza del ruido coloreado fraccionario es

$$\mathbb{E}(\varepsilon_l^2) = \sigma^2 2^{2H} (\delta_n)^{2H+d+\alpha}. \tag{2.21}$$

□

2.4. Covariables correlacionadas

Respecto a las covariables X_j , para $j = 1, \dots, p$, en el modelo de regresión 2.1, asumiremos que son centradas y correlacionadas localmente, donde su función de covarianza es $\chi(z, z') = (\chi_{jk}(z, z'))_{j,k=1:p}$, con

$$\chi_{jk}(z, z') = \mathbb{E}(X_j(z)X_k(z')), \tag{2.22}$$

es decir, 2.22 muestra la correlación local entre las observación z y z' a través de las covariables X_{jth} -ésima y X_{kth} -ésima.

Ahora bien, bajo la consideración anterior se planten los siguientes supuestos:

Hipótesis 2.4. *Se tiene que*

- i) *La función de covarianza χ es simétrica positiva definida.*
- ii) *χ is α_χ -Hölder continua; Es decir, existe una constante $C_\chi > 0$ tal que*

$$|\chi(z, z') - \chi(z_l, z'_l)| \leq C_\chi (\|z - z_l\| + \|z' - z'_l\|)^{\alpha_\chi}.$$

Por otra parte, consideremos la función $\Gamma(z, z') = (\Gamma_{jk}(z, z'))_{j,k=1:n}$ definida por

$$\Gamma_{jk}(z, z') = \text{Cov}(X_j(z)X_k(z), X_j(z')X_k(z')). \quad (2.23)$$

Hipótesis 2.5. Γ satisface las siguientes condiciones:

i) La función de covarianza Γ es simétrica y positiva definida.

ii) Γ es α_Γ -Hölder continua, es decir, existe una constante $C_\Gamma > 0$ tal que

$$|\Gamma(z, z') - \Gamma(z_l, z'_l)| \leq C_\Gamma(\|z - z_l\| + \|z' - z'_l\|)^{\alpha_\Gamma}.$$

iii) Más aún, Γ es tal que

$$|\Gamma(z, z')| \leq C_{k,d,\theta} \delta^{d+1+\theta},$$

para $z, z' \in \mathbb{R}^{d+1}$ tal que $\|z - z'\| > k\delta$, para $\delta > 0$ y algún $k \in \mathbb{N}$, $-(d+1) < \theta < d+1$, y $C_{k,d,\theta} \geq 0$.

Observación 2.8. A partir del supuesto 2.4 se consideraron dos tipos de interacciones entre las covariables X'_j s:

- **Interacción débil:** cuando el parámetro $\theta \leq 0$. Por ejemplo, el caso de independencia entre las covariables se obtiene cuando $\theta = 0$, el caso de las covariables k -dependientes corresponde cuando $C_{k,d,\theta} = 0$.
- **Interacción fuerte:** cuando el parámetro $\theta > 0$ entonces la densidad espectral de la función de covarianza Γ es singular en cero, así Γ tiene cola pesada. La dependencia fraccionaria en tiempo corresponde para $\theta = 2H - 1$ y esto tendrá larga dependencia cuando $\theta > 0$. Es decir, si $H > \frac{1}{2}$, la dependencia espacial fraccionaria coloreada corresponde para $\theta = 2H - 1 + \alpha$.

2.5. Estimador de mínimos cuadrados ponderados

Considerando un conjunto de datos dados, los parámetros locales del modelo de regresión ponderado en 2.1 son estimados usando el procedimiento de mínimos cuadrados ponderados.

Sea $\beta(z_i)$ el vector de parámetros locales para la observación espacio-temporal z_i , dado por:

$$\beta(z_i) = (\beta_0(z_i), \beta_1(z_i), \dots, \beta_p(z_i))^T. \quad (2.24)$$

El supra índice T representa la transpuesta de la matriz o vector transpuesto de parámetros locales.

Ahora bien, los parámetros locales $\beta(z_i)$ en el punto z_i son estimados por la siguiente expresión:

$$\hat{\beta}(z_i) = [X^T \mathcal{W}(z_i) X]^{-1} X^T \mathcal{W}(z_i) Y, \quad (2.25)$$

donde X es una matriz de orden $n \times (p + 1)$ correspondiente a las entradas de las covariables, Y es el vector n -dimensional de observaciones espacio-temporal y $\mathcal{W}(z_i)$ es una matriz de orden $n \times n$ conocida como matriz de pesos, que tiene la siguiente forma:

$$\mathcal{W}(z_i) = \text{diag}(\mathcal{W}_{i1}, \dots, \mathcal{W}_{in}). \quad (2.26)$$

Consideremos también, una función kernel $K : \mathbb{R}^{d+1} \rightarrow \mathbb{R}$, simétrica y positiva definida, tal que

$$\int_{\mathbb{R}^{d+1}} K(z) dz = 1.$$

Por otro lado, los pesos \mathcal{W}_{ij} , para $j = 1, \dots, n$, son obtenidos usualmente a través de la función kernel K , quien es adaptativa en términos de la proximidad que hay entre los puntos de la data y el punto de observación z_i . Es decir,

$$\mathcal{W}_{il} = K_h(z_l - z_i), \quad (2.27)$$

siendo $K_h(z) = K(z/h)$.

Observación 2.9. Aquí, h es un parámetro que no toma valores negativos y se conoce como banda de ajuste entre las observaciones. Dicho h produce un decaimiento de influencia entre las distancias. Esto es: Las observaciones z_l cercanas a z_i tienen la mayor influencia sobre las estimaciones de los parámetros locales en el punto z_i .

Adicionalmente, consideremos que la función kernel K satisface el siguiente supuesto:

Hipótesis 2.6. Consideremos una función Kernel K , que cumple las siguientes condiciones:

i) La función kernel K es acotada, es decir, $\|K\|_{\infty} < \infty$.

ii) K es α_K -Hölder continua, es decir, existe $C_K > 0$ tal que

$$|K(z) - K(z')| < C_K \|z - z'\|^{\alpha_K}.$$

iii) $\int_{\mathbb{R}^{d+1}} \max(\|z\|^{\alpha_K}, \|z\|^{\alpha_{\chi}}, \|z\|^{\alpha_{\Gamma}}) K(z) dz < \infty$

Observación 2.10. Note que bajo el supuesto 2.6, la función kernel K satisface $\int_{\mathbb{R}^{d+1}} zK(z)dz = 0$. Además, la función kernel adaptativa que mayormente se usa para el cálculo de los W_{il} , es la función kernel Gaussiana $K(z) = \frac{1}{\sqrt{2\pi}} e^{-(d^{s,t}(z))^2/2}$, en donde la distancia espacio-temporal $d^{s,t}$ es obtenida en función de la distancia temporal $d^t = |t|$ y la distancia espacial $d^u = \|u\|$; Por ejemplo, $(d^{s,t}(z))^2 = \mu^t (d^t)^2 + \mu^s (d^u)^2$ en donde μ^t y μ^s son factores de escala temporal y espacial respectivamente.

2.6. Consistencia del estimador

En esta sección estudiaremos la consistencia del estimador local de mínimos cuadrados ponderados $\hat{\beta}(z_i)$ en (2.25) desde 2.1. En este sentido, si sustituimos $Y = X\beta(z_i) + \varepsilon$ en (2.25) obtenemos que:

$$\begin{aligned} \hat{\beta}(z_i) &= (X^T \mathcal{W}(z_i) X)^{-1} X^T \mathcal{W}(z_i) (X\beta(z_i) + \varepsilon) \\ &= (X^T \mathcal{W}(z_i) X)^{-1} (X^T \mathcal{W}(z_i) X) \beta(z_i) + (X^T \mathcal{W}(z_i) X)^{-1} X^T \mathcal{W}(z_i) \varepsilon \\ &= \beta(z_i) + (X^T \mathcal{W}(z_i) X)^{-1} X^T \mathcal{W}(z_i) \varepsilon. \end{aligned}$$

Entonces,

$$\mathbb{E}(\hat{\beta}(z_i)) = \beta(z_i) + \mathbb{E}(X^T \mathcal{W}(z_i) X)^{-1} X^T \mathcal{W}(z_i) \mathbb{E}(\varepsilon|X) = \beta(z_i),$$

de donde asumiremos 2.1 y se tiene que $\mathbb{E}(\varepsilon|X) = \mathbb{E}(\varepsilon) = 0$. Así, el estimador $\hat{\beta}(z_i)$ es insesgado, y el error de estimación se escribe como:

$$\hat{\beta}(z_i) - \beta(z_i) = (X^T \mathcal{W}(z_i) X)^{-1} (X^T \mathcal{W}(z_i) \varepsilon). \quad (2.28)$$

Proposición 2.3. Bajo los supuestos 2.4, 2.5 y 2.6, para $\theta > 0$, tenemos que

$$\frac{1}{nh^{d+1}} (X^T \mathcal{W}(z_i) X) \xrightarrow[n \rightarrow \infty]{a.s.} \chi(z_i, z_i) = \mathbb{E}[X^T \mathcal{W}(z_i) X].$$

La convergencia en probabilidad se asegura para $d + 1 + \theta > 0$.

Demostración. La jk_{th} componente de la matriz $X^T \mathcal{W}(z_i) X$ es

$$(X^T \mathcal{W}(z_i) X)_{jk} = \sum_{l=1}^n X_{lj} X_{lk} \mathcal{W}_{il}. \quad (2.29)$$

Definimos $\chi_{n,h} \simeq \chi_h$ y definimos $\Gamma_{n,h} \simeq \Gamma_h$ de manera que el $\lim_{h \rightarrow 0} \lim_{n \rightarrow \infty} \chi_{n,h} = \lim_{h \rightarrow 0} \chi_h$ y el $\lim_{h \rightarrow 0} \lim_{n \rightarrow \infty} \Gamma_{n,h} = \lim_{h \rightarrow 0} \Gamma_h$, respectivamente. Es decir, para un n lo

suficientemente grande y h suficientemente pequeño, $\chi_{n,h}$ y $\Gamma_{n,h}$ son aproximadamente igual a χ_h y Γ_h respectivamente. De forma análoga definimos $\Gamma_{n,h} \preceq \Gamma_h$ para denotar que el $\lim_{h \rightarrow 0} \lim_{n \rightarrow \infty} \Gamma_{n,h} \leq \lim_{h \rightarrow 0} \Gamma_h$. En particular escribimos $\Gamma_{n,h} \simeq \Gamma$ para indicar que Γ es el límite de una sucesión $\Gamma_{n,h}$, para un n suficientemente grande y un h lo suficientemente pequeño. Esto último, a modo tener una mejor interpretación de las aproximaciones límites para el resto de las demostraciones dentro del capítulo.

Ahora bien, estudiamos la esperanza asintótica de (2.29), bajo la hipótesis 2.4 obteniendo

$$\begin{aligned}
\frac{1}{nh^{d+1}} \mathbb{E} \left((X^T \mathcal{W}(z_i) X)_{jk} \right) &= \frac{1}{nh^{d+1}} \sum_{l=1}^n \mathbb{E} (X_{lj} X_{lk}) K_h(z_l - z_i) \\
&= \frac{1}{h^{d+1}} \int_{\mathbb{R}^{d+1}} \sum_{l=1}^n \chi_{jk}(z_l, z_l) K_h(z_l - z_i) \mathbb{I}_{V(z_l)}(z) dz \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{1}{h^{d+1}} \int_{\mathbb{R}^{d+1}} \chi_{jk}(z, z) K_h(z - z_i) dz \\
&= \int_{\mathbb{R}^{d+1}} \chi_{jk}(z_i + hz, z_i + hz) K(z) dz \\
&\stackrel{h \rightarrow 0}{\simeq} \chi_{jk}(z_i, z_i) \int_{\mathbb{R}^{d+1}} K(z) dz + C_\chi |h|^{\alpha_\chi} \int_{\mathbb{R}^{d+1}} \|z\|^{\alpha_\chi} K(z) dz \\
&= \chi_{jk}(z_i, z_i) + \mathcal{O}(|h|^{\alpha_\chi})
\end{aligned} \tag{2.30}$$

Note que la condición 2.5 implica que la matriz de covarianza $\chi(z_i, z_i) = (\chi_{jk}(z_i, z_i))_{j,k=1:n}$ es una matriz invertible.

Ahora, calculamos la varianza de (2.29).

$$\begin{aligned}
\text{Var} \left((X^T \mathcal{W}(z_i) X)_{jk} \right) &= \sum_{l, l'=1}^n \Gamma_{jk}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \\
&= \sum_{l=1}^n \Gamma_{jk}(z_l, z_l) \mathcal{W}_{il}^2 \\
&+ \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq k\delta_n}} \Gamma_{jk}(z_l, z_{l'}) \mathcal{W}_{il} \mathcal{W}_{il'} \\
&+ \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| > k\delta_n}} \Gamma_{jk}(z_l, z_{l'}) \mathcal{W}_{il} \mathcal{W}_{il'} \\
&:= D_{jk,n}^{(1)}(z_i) + D_{jk,n}^{(2)}(z_i) + D_{jk,n}^{(3)}(z_i),
\end{aligned} \tag{2.31}$$

Primero, estudiamos el término $D_{jk,n}^{(1)}(z_i)$ en (2.31). Bajo la hipótesis de regularidad 2.6.

$$\begin{aligned}
\frac{1}{nh^{d+1}} D_{jk,n}^{(1)}(z_i) &= \frac{1}{nh^{d+1}} \sum_{l=1}^n \Gamma_{jk}(z_l, z_l) K_h^2(z_l - z_i) \\
&= \int_{\mathbb{R}^{d+1}} \sum_{l=1}^n \Gamma_{jk}(z_l, z_l) \frac{1}{h^{d+1}} K_h^2(z_l - z_i) \mathbb{I}_{V(z_l)}(z) dz \\
&\stackrel{n \rightarrow \infty}{\approx} \int_{\mathbb{R}^{d+1}} \Gamma_{jk}(z, z) \frac{1}{h^{d+1}} K_h^2(z - z_i) dz \\
&= \int_{\mathbb{R}^{d+1}} \Gamma_{jk}(z_i + hz, z_i + hz) K^2(z) dz \\
&\stackrel{h \rightarrow 0}{\approx} \Gamma_{jk}(z_i, z_i) \|K\|_2^2 + C_\Gamma |h|^{\alpha_\Gamma} \int_{\mathbb{R}^{d+1}} \|z\|^{\alpha_\Gamma} K^2(z) dz \\
&= \Gamma_{jk}(z_i, z_i) \|K\|_2^2 + \mathcal{O}(|h|^{\alpha_\Gamma})
\end{aligned} \tag{2.32}$$

Segundo, consideramos el término $D_{jk,n}^{(2)}(z_i)$ en (2.31), es decir, cuando $0 < \|z_l - z_{l'}\| \leq k\delta_n$

$$\begin{aligned}
\frac{1}{nh^{d+1}} D_{jk,n}^{(2)}(z_i) &= \frac{1}{nh^{d+1}} \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq k\delta_n}} \Gamma_{jk}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \\
&= \frac{1}{nh^{d+1}} \left[\sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq k\delta_n}} \Gamma_{jk}(z_l, z_l) K_h(z_l - z_i) K_h(z_l - z_i) \right. \\
&+ \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq k\delta_n}} \Gamma_{jk}(z_l, z_l) K_h(z_l - z_i) (K_h(z_{l'} - z_i) - K_h(z_l - z_i)) \\
&+ \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq k\delta_n}} (\Gamma_{jk}(z_l, z_{l'}) - \Gamma_{jk}(z_l, z_l)) K_h(z_l - z_i) K_h(z_l - z_i) \\
&\left. + \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq k\delta_n}} (\Gamma_{jk}(z_l, z_{l'}) - \Gamma_{jk}(z_l, z_l)) K_h(z_l - z_i) (K_h(z_{l'} - z_i) - K_h(z_l - z_i)) \right]
\end{aligned} \tag{2.33}$$

Bajo hipótesis de regularidad en 2.5 y 2.6

$$\begin{aligned}
\frac{1}{nh^{d+1}} D_{jk,n}^{(2)}(z_i) &\leq \frac{1}{nh^{d+1}} \left[\sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq k\delta_n}} \Gamma_{jk}(z_l, z_l) K_h^2(z_l - z_i) \right. \\
&+ C_K \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq k\delta_n}} \Gamma_{jk}(z_l, z_l) K_h(z_l - z_i) \|z_l - z_{l'}\|^{\alpha_K} \\
&+ C_\Gamma \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq k\delta_n}} \|z_l - z_{l'}\|^{\alpha_\Gamma} K_h^2(z_l - z_i) \\
&\left. + C_k C_\Gamma \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq k\delta_n}} \|z_l - z_{l'}\|^{\alpha_\Gamma} K_h(z_l - z_i) \|z_l - z_{l'}\|^{\alpha_K} \right] \\
&= D_{jk,n}^{(2,1)} + D_{jk,n}^{(2,2)} + D_{jk,n}^{(2,3)} + D_{jk,n}^{(2,4)}, \tag{2.34}
\end{aligned}$$

donde C_k y C_Γ son constante positivas que se resultan de la regularidad de K y Γ que se asume desde las hipótesis 2.5 y 2.6 respectivamente.

Note que de manera similar desde (2.48) tenemos que $\frac{1}{n} \sum_{l'=1}^n \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq k\delta_n\}} \stackrel{n \rightarrow \infty}{\simeq} \frac{k^{d+1}}{n}$. Entonces, desde 2.34, hipótesis 2.5 y 2.6

$$\begin{aligned}
D_{jk,n}^{(2,1)} &= \frac{1}{nh^{d+1}} \sum_{l=1}^n \Gamma_{jk}(z_l, z_l) K_h^2(z_l - z_i) \sum_{l'=1}^n \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq k\delta_n\}} \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{k^{d+1}}{h^{d+1}} \int_{\mathbb{R}^{d+1}} \Gamma_{jk}(z, z) K_h^2(z - z_i) dz \\
&= k^{d+1} \int_{\mathbb{R}^{d+1}} \Gamma_{jk}(z_i + hz, z_i + hz) K^2(z) dz \\
&\stackrel{h \rightarrow 0}{\simeq} k^{d+1} \Gamma_{jk}(z_i, z_i) \int_{\mathbb{R}^{d+1}} K^2(z) dz \\
&+ k^{d+1} C_\Gamma |h|^{\alpha_\Gamma} \int_{\mathbb{R}^{d+1}} \|z\|^{\alpha_\Gamma} K(z) dz \\
&= k^{d+1} \Gamma_{jk}(z_i, z_i) \|K\|_2^2 + \mathcal{O}(|h|^{\alpha_\Gamma}). \tag{2.35}
\end{aligned}$$

Nuevamente desde 2.34 y bajo hipótesis 2.5 y 2.6

$$\begin{aligned}
\frac{1}{(\delta_n)^{\alpha_k}} D_{jk,n}^{(2,2)} &= \frac{C_K}{nh^{d+1}(\delta_n)^{\alpha_k}} \sum_{l=1}^n \Gamma_{jk}(z_l, z_l) K_h(z_l - z_i) \sum_{l'=1}^n \|z_l - z_{l'}\|^{\alpha_k} \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq k\delta_n\}} \\
&\leq \frac{C_K(k\delta_n)^{\alpha_k}}{nh^{d+1}(\delta_n)^{\alpha_k}} \sum_{l=1}^n \Gamma_{jk}(z_l, z_l) K_h(z_l - z_i) \sum_{l'=1}^n \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq k\delta_n\}} \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{k^{\alpha_k+d+1} C_K}{h^{d+1}} \int_{\mathbb{R}^{d+1}} \Gamma_{jk}(z, z) K_h(z - z_i) dz \\
&= k^{\alpha_k+d+1} C_K \int_{\mathbb{R}^{d+1}} \Gamma_{jk}(z_i + hz, z_i + hz) K(z) dz \\
&\stackrel{h \rightarrow 0}{\simeq} k^{\alpha_k+d+1} C_K \Gamma_{jk}(z_i, z_i) \int_{\mathbb{R}^{d+1}} K(z) dz \\
&\quad + k^{\alpha_k+d+1} C_K C_\Gamma |h|^{\alpha_\Gamma} \int_{\mathbb{R}^{d+1}} \|z\|^{\alpha_\Gamma} K(z) dz \\
&= k^{\alpha_k+d+1} C_K \Gamma_{jk}(z_i, z_i) + \mathcal{O}(|h|^{\alpha_\Gamma}).
\end{aligned} \tag{2.36}$$

Ahora, para el siguiente término a considerar en 2.34 y la hipótesis 2.5

$$\begin{aligned}
\frac{1}{(\delta_n)^{\alpha_\Gamma}} D_{jk,n}^{(2,3)} &= \frac{C_\Gamma}{nh^{d+1}(\delta_n)^{\alpha_\Gamma}} \sum_{l=1}^n K_h^2(z_l - z_i) \sum_{l'=1}^n \|z_l - z_{l'}\|^{\alpha_\Gamma} \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq k\delta_n\}} \\
&\leq \frac{C_\Gamma(3\delta_n)^{\alpha_\Gamma}}{nh^{d+1}(\delta_n)^{\alpha_\Gamma}} \sum_{l=1}^n K_h^2(z_l - z_i) \sum_{l'=1}^n \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq k\delta_n\}} \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{k^{\alpha_\Gamma+d+1} C_\Gamma}{h^{d+1}} \int_{\mathbb{R}^{d+1}} K_h^2(z - z_i) dz \\
&= k^{\alpha_\Gamma+d+1} C_\Gamma \|K\|_2^2.
\end{aligned} \tag{2.37}$$

Luego, como último término de 2.34 y considerando la regularidad desde las hipótesis 2.5 y 2.6

$$\begin{aligned}
\frac{1}{(\delta_n)^{\alpha_\Gamma+\alpha_k}} D_{jk,n}^{(2,4)} &= \frac{C_K C_\Gamma}{nh^{d+1}(\delta_n)^{\alpha_\Gamma+\alpha_k}} \sum_{l=1}^n K_h(z_l - z_i) \sum_{l'=1}^n \|z_l - z_{l'}\|^{\alpha_\Gamma+\alpha_k} \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq k\delta_n\}} \\
&\leq \frac{C_K C_\Gamma (k\delta_n)^{\alpha_\Gamma+\alpha_k}}{nh^{d+1}(\delta_n)^{\alpha_\Gamma+\alpha_k}} \sum_{l=1}^n K_h(z_l - z_i) \sum_{l'=1}^n \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq k\delta_n\}} \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{k^{\alpha_\Gamma+\alpha_k+d+1} C_K C_\Gamma}{h^{d+1}} \int_{\mathbb{R}^{d+1}} K_h(z - z_i) dz \\
&= k^{\alpha_\Gamma+\alpha_k+d+1} C_K C_\Gamma.
\end{aligned} \tag{2.38}$$

Así, desde (2.35), (2.36), (2.37), y (2.38) tenemos

$$\begin{aligned} & \frac{1}{nh^{d+1}} D_{jk,n}^{(2)}(z_i) \\ & \stackrel{\leq}{\underset{n \rightarrow \infty, h \rightarrow 0}{\leq}} k^{d+1} \Gamma_{jk}(z_i, z_i) \|K\|_2^2 + \mathcal{O}(|h|^{\alpha_\Gamma} \vee (\delta_n)^{\alpha_K} \vee (\delta_n)^{\alpha_\Gamma}). \end{aligned} \quad (2.39)$$

Finalmente consideremos el termino $D_{jk,n}^{(3)}(z_i)$, usando la hipótesis 2.5 obtenemos

$$\begin{aligned} & \frac{1}{n^2 h^{2(d+1)}} D_{jk,n}^{(3)} \\ & = \frac{1}{n^2 h^{2(d+1)}} \sum_{1 \leq l \neq l' \leq n} \Gamma_{jk}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \mathbb{I}_{\{\|z_l - z_{l'}\| > k\delta_n\}} \\ & \leq \frac{C_{k,d,\beta} (\delta_n)^{d+1+\beta}}{n^2 h^{2(d+1)}} \sum_{1 \leq l \neq l' \leq n} K_h(z_l - z_i) K_h(z_{l'} - z_i) \\ & \stackrel{\simeq}{\underset{n \rightarrow \infty}{\simeq}} \frac{C_{k,d,\beta} (\delta_n)^{d+1+\beta}}{h^{2(d+1)}} \int_{\mathbb{R}^{d+1}} \int_{\mathbb{R}^{d+1}} K_h(z - z_i) K_h(z' - z_i) dz dz' \\ & = C_{k,d,\beta} (\delta_n)^{d+1+\beta}. \end{aligned} \quad (2.40)$$

Donde $C_{k,d,\beta}$ es una constante positiva que resulta de asumir el punto *iii*) de la hipótesis 2.5. Luego, sustituyendo (2.32), (2.39) and (2.40) en la ecuación (2.31), y usando que $2\lambda(\mathcal{S}^{d-1})(\delta_n)^{d+1} = 1/n$ obtenemos

$$\frac{1}{n^2 h^{2(d+1)}} \text{Var} \left((X^T \mathcal{W}(z_i) X)_{jk} \right) \leq \frac{(1 + k^{d+1}) \Gamma_{jk}(z_i, z_i) \|K\|_2^2}{nh^{d+1}} + C_{k,d,\beta} (\delta_n)^{d+1+\beta} \simeq \frac{C}{n^{1+\theta}},$$

donde $\theta = \frac{\beta}{d+1} > 0$ y donde $\beta > 0$; es decir, el L^2 radio de $\frac{1}{nh^{d+1}} (X^T \mathcal{W}(z_i) X)_{jk}$ es más rápido que $1/n$. Entonces, el lema de Borell-Cantelli nos permitió obtener

$$\frac{1}{nh^{d+1}} (X^T \mathcal{W}(z_i) X)_{jk} \xrightarrow[n \rightarrow \infty]{a.s.} \chi_{j,k}(z_i, z_i).$$

□

Teorema 2.1. *Si el modelo de regresión 2.1 satisface del supuesto 2.1 al 2.6, entonces, el estimador local de mínimos cuadrados ponderados $\hat{\beta}(z_i)$ obtenido en 2.25 es fuertemente consistente para $2H + \alpha > 1$ y $\theta > 0$. Es decir:*

$$\left| \hat{\beta}(z_i) - \beta(z_i) \right| \xrightarrow[n \rightarrow \infty]{a.s.} 0.$$

Además, la convergencia en probabilidad se asegura para $2H + d + \alpha > 0$ y $d + 1 + \theta > 0$.

Demostración. Para estudiar la consistencia del estimador $\hat{\beta}(z_i)$, probaremos que existe una sucesión de normalización apropiada $(b_n)_{n \geq 1}$ de constantes positivas con $b_n \rightarrow \infty$ cuando $n \rightarrow \infty$, y tal que

$$i) \quad b_n^{-1}(X^T \mathcal{W}(z_i)X) \rightarrow \chi(z_i, z_i) = \mathbb{E}[X^T \mathcal{W}(z_i)X], \text{ as } n \rightarrow +\infty.$$

$$ii) \quad b_n^{-1}(X^T \mathcal{W}(z_i)\varepsilon) \rightarrow 0, n \rightarrow +\infty.$$

Para la demostración de *i)* citaremos la proposición 2.3 relacionado con la convergencia casi segura del término $(X^T \mathcal{W}(z_i)X)$ en (2.28).

Por tanto, por la proposición 2.3, queda estudiar el comportamiento asintótico de $X^T \mathcal{W}(z_i)\varepsilon$ cuando

$n \rightarrow \infty$. El componente j_{th} de $X^T \mathcal{W}(z_i)\varepsilon$ es

$$(X^T \mathcal{W}(z_i)\varepsilon)_j = \sum_{l=1}^n X_{lj} \mathcal{W}_{il} \varepsilon_l. \quad (2.41)$$

A partir de la hipótesis 2.1, se obtiene $\mathbb{E}\left((X^T \mathcal{W}(z_i)\varepsilon)_j\right) = 0$. Cuando la varianza de $(X^T \mathcal{W}(z_i)\varepsilon)_j$,

$$\begin{aligned} \text{Var}\left((X^T \mathcal{W}(z_i)\varepsilon)_j\right) &= \mathbb{E}\left((X^T \mathcal{W}(z_i)\varepsilon)_j^2\right) = \sum_{l, l'=1}^n \chi_{jj}(z_l, z_{l'}) \mathcal{W}_{il} \mathcal{W}_{il'} \mathbb{E}(\varepsilon_l \varepsilon_{l'}) \\ &= \sum_{l=1}^n \chi_{jj}(z_l, z_l) \mathcal{W}_{il}^2 \mathbb{E}(\varepsilon_l^2) \\ &\quad + \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} \chi_{jj}(z_l, z_{l'}) \mathcal{W}_{il} \mathcal{W}_{il'} \mathbb{E}(\varepsilon_l \varepsilon_{l'}) \\ &\quad + \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| > 3\delta_n}} \chi_{jj}(z_l, z_{l'}) \mathcal{W}_{il} \mathcal{W}_{il'} \mathbb{E}(\varepsilon_l \varepsilon_{l'}) \\ &:= A_{j,n}^{(1)}(z_i) + A_{j,n}^{(2)}(z_i) + A_{j,n}^{(3)}(z_i), \quad (2.42) \end{aligned}$$

donde partimos la suma en tres términos asociados a la distancia entre las observaciones z_l y $z_{l'}$.

Primero, estudiaremos el término $A_{j,n}^{(1)}(z_i)$ en 2.42. A partir de la condición de regularidad 2.4

$$\begin{aligned}
& \frac{1}{nh^{d+1}(\delta_n)^{2H+d+\alpha}} A_{j,n}^{(1)}(z_i) \\
&= \frac{1}{nh^{d+1}(\delta_n)^{2H+d+\alpha}} \sum_{l=1}^n \chi_{jj}(z_l, z_l) \mathcal{W}_{il}^2 \mathbb{E}(\varepsilon_l^2) \\
&= \frac{2^{2H} \sigma^2}{nh^{d+1}} \sum_{l=1}^n \chi_{jj}(z_l, z_l) K_h^2(z_l - z_i) \\
&= 2^{2H} \sigma^2 \int_{\mathbb{R}^{d+1}} \sum_{l=1}^n \chi_{jj}(z_l, z_l) \frac{1}{h^{d+1}} K_h^2(z_l - z_i) \mathbb{1}_{V(z_l)}(z) dz \\
&\stackrel{n \rightarrow \infty}{\approx} 2^{2H} \sigma^2 \int_{\mathbb{R}^{d+1}} \chi_{jj}(z, z) \frac{1}{h^{d+1}} K_h^2(z - z_i) dz \\
&= 2^{2H} \sigma^2 \int_{\mathbb{R}^{d+1}} \chi_{jj}(z_i + hz, z_i + hz) K^2(z) dz \\
&\stackrel{h \rightarrow 0}{\approx} 2^{2H} \sigma^2 \chi_{jj}(z_i, z_i) \|K\|_2^2 + |h|^{\alpha_x} 2^{2H} \sigma^2 \int_{\mathbb{R}^{d+1}} \|z\|^{\alpha_x} K^2(z) dz \\
&= 2^{2H} \sigma^2 \chi_{jj}(z_i, z_i) \|K\|_2^2 + \mathcal{O}(|h|^{\alpha_x}) \\
&= C_1(H) \chi_{jj}(z_i, z_i) + \mathcal{O}(|h|^{\alpha_x}),
\end{aligned} \tag{2.43}$$

donde $C_1(H) = 2^{2H} \sigma^2 \|K\|_2^2$.

Como segundo paso, consideramos el término $A_{j,n}^{(2)}(z_i)$ in (2.42), es decir, $0 < \|z_l - z_{l'}\| \leq 3\delta_n$

$$\begin{aligned}
& \frac{1}{nh^{d+1}(\delta_n)^{2H+d+\alpha}} A_{j,n}^{(2)}(z_i) \\
&= \frac{1}{nh^{d+1}(\delta_n)^{2H+d+\alpha}} \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} \chi_{jj}(z_l, z_{l'}) \mathcal{W}_{il} \mathcal{W}_{il'} \mathbb{E}(\varepsilon_l \varepsilon_{l'}) \\
&= \frac{1}{nh^{d+1}(\delta_n)^{2H+d+\alpha}} \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \mathbb{E}(\varepsilon_l \varepsilon_{l'})
\end{aligned} \tag{2.44}$$

Podemos acotar los términos de la covarianza $\mathbb{E}(\varepsilon_l \varepsilon_{l'})$ cuando $0 < \|z_l - z_{l'}\| \leq 3\delta_n$ por

$$\begin{aligned}
\mathbb{E}(\varepsilon_l \varepsilon_{l'}) &= \frac{1}{2} (|t_l - t_{l'} + 2\delta_n|^{2H} + |t_l - t_{l'} - 2\delta_n|^{2H} - 2|t_l - t_{l'}|^{2H}) (\delta_n)^{d+\alpha} \text{Cov} \left(\mathbb{I}_{\{\|u - u_l/\delta_n\| \leq 1\}}, \mathbb{I}_{\{\|u - u_{l'}/\delta_n\| \leq 1\}} \right) \\
&\leq (2\delta_n)^{2H} (\delta_n)^{d+\alpha} \text{Var}^{1/2}(\mathbb{I}_{\{\|u - u_l/\delta_n\| \leq 1\}}) \text{Var}^{1/2}(\mathbb{I}_{\{\|u - u_{l'}/\delta_n\| \leq 1\}}) \\
&\leq 2^{2H} (\delta_n)^{2H+d+\alpha} \text{Var}(\mathbb{I}_{\|u\| \leq 1}) \\
&= 2^{2H} \sigma^2 (\delta_n)^{2H+d+\alpha}.
\end{aligned} \tag{2.45}$$

Introduciendo la desigualdad (2.45) en la ecuación (2.44) se obtiene

$$\begin{aligned}
&\frac{1}{nh^{d+1} (\delta_n)^{2H+d+\alpha}} A_{j,n}^{(2)}(z_i) \\
&\leq \frac{2^{2H} \sigma^2}{nh^{d+1}} \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \\
&= \frac{2^{2H} \sigma^2}{nh^{d+1}} \left[\sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} \chi_{jj}(z_l, z_l) K_h(z_l - z_i) K_h(z_l - z_i) \right. \\
&+ \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} \chi_{jj}(z_l, z_l) K_h(z_l - z_i) (K_h(z_{l'} - z_i) - K_h(z_l - z_i)) \\
&+ \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} (\chi_{jj}(z_l, z_{l'}) - \chi_{jj}(z_l, z_l)) K_h(z_l - z_i) K_h(z_l - z_i) \\
&+ \left. \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} (\chi_{jj}(z_l, z_{l'}) - \chi_{jj}(z_l, z_l)) K_h(z_l - z_i) (K_h(z_{l'} - z_i) - K_h(z_l - z_i)) \right]
\end{aligned} \tag{2.46}$$

A partir del supuesto 2.4 y 2.44

$$\begin{aligned}
& \frac{1}{nh^{d+1}(\delta_n)^{2H+d+\alpha}} A_{j,n}^{(2)}(z_i) \\
& \leq \frac{2^{2H}\sigma^2}{nh^{d+1}} \left[\sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} \chi_{jj}(z_l, z_l) K_h^2(z_l - z_i) \right. \\
& + C_K \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} \chi_{jj}(z_l, z_l) K_h(z_l - z_i) \|z_l - z_{l'}\|^{\alpha_K} \\
& + C_\chi \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} \|z_l - z_{l'}\|^{\alpha_\chi} K_h^2(z_l - z_i) \\
& \left. + C_k C_\chi \sum_{\substack{1 \leq l \neq l' \leq n \\ \|z_l - z_{l'}\| \leq 3\delta_n}} \|z_l - z_{l'}\|^{\alpha_\chi} K_h(z_l - z_i) \|z_l - z_{l'}\|^{\alpha_K} \right] \\
& = A_{j,n}^{(2,1)} + A_{j,n}^{(2,2)} + A_{j,n}^{(2,3)} + A_{j,n}^{(2,4)}. \tag{2.47}
\end{aligned}$$

Note que

$$\frac{1}{n} \sum_{l'=1}^n \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq 3\delta_n\}} \stackrel{n \rightarrow \infty}{\simeq} 3^{d+1} \int_{\mathbb{R}^{d+1}} \mathbb{I}_{V(z_l)}(z') dz' = 3^{d+1} \lambda(V(z_l)) = \frac{3^{d+1}}{n}. \tag{2.48}$$

Entonces, a partir de (2.47), (2.48) y bajo condiciones de regularidad asumidas desde 2.4

$$\begin{aligned}
A_{j,n}^{(2,1)} &= \frac{2^{2H}\sigma^2}{nh^{d+1}} \sum_{l=1}^n \chi_{jj}(z_l, z_l) K_h^2(z_l - z_i) \sum_{l'=1}^n \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq 3\delta_n\}} \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{2^{2H}3^{d+1}\sigma^2}{h^{d+1}} \int_{\mathbb{R}^{d+1}} \chi_{jj}(z, z) K_h^2(z - z_i) dz \\
&= 2^{2H}3^{d+1}\sigma^2 \int_{\mathbb{R}^{d+1}} \chi_{jj}(z_i + hz, z_i + hz) K^2(z) dz \\
&\stackrel{h \rightarrow 0}{\simeq} 2^{2H}3^{d+1}\sigma^2 \chi_{jj}(z_i, z_i) \int_{\mathbb{R}^{d+1}} K^2(z) dz \\
&+ 2^{2H}3^{d+1}\sigma^2 C_\chi |h|^{\alpha_\chi} \int_{\mathbb{R}^{d+1}} \|z\|^{\alpha_\chi} K(z) dz \\
&= 2^{2H}3^{d+1}\sigma^2 \chi_{jj}(z_i, z_i) \|K\|_2^2 + \mathcal{O}(|h|^{\alpha_\chi}). \tag{2.49}
\end{aligned}$$

Ahora bien, para el segundo término de 2.47, usando 2.48, y bajo las hipótesis 2.4 y 2.6

$$\begin{aligned}
\frac{1}{(\delta_n)^{\alpha_k}} A_{j,n}^{(2,2)} &= \frac{2^{2H} \sigma^2 C_K}{nh^{d+1} (\delta_n)^{\alpha_k}} \sum_{l=1}^n \chi_{jj}(z_l, z_l) K_h(z_l - z_i) \sum_{l'=1}^n \|z_l - z_{l'}\|^{\alpha_k} \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq 3\delta_n\}} \\
&\leq \frac{2^{2H} \sigma^2 C_K (3\delta_n)^{\alpha_k}}{nh^{d+1} (\delta_n)^{\alpha_k}} \sum_{l=1}^n \chi_{jj}(z_l, z_l) K_h(z_l - z_i) \sum_{l'=1}^n \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq 3\delta_n\}} \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{2^{2H} 3^{\alpha_k+d+1} \sigma^2 C_K}{h^{d+1}} \int_{\mathbb{R}^{d+1}} \chi_{jj}(z, z) K_h(z - z_i) dz \\
&= 2^{2H} 3^{\alpha_k+d+1} \sigma^2 C_K \int_{\mathbb{R}^{d+1}} \chi_{jj}(z_i + hz, z_i + hz) K(z) dz \\
&\stackrel{h \rightarrow 0}{\simeq} 2^{2H} 3^{\alpha_k+d+1} \sigma^2 C_K \chi_{jj}(z_i, z_i) \int_{\mathbb{R}^{d+1}} K(z) dz \\
&\quad + 2^{2H} 3^{\alpha_k+d+1} \sigma^2 C_K C_\chi |h|^{\alpha_\chi} \int_{\mathbb{R}^{d+1}} \|z\|^{\alpha_\chi} K(z) dz \\
&= 2^{2H} 3^{\alpha_k+d+1} \sigma^2 C_K \chi_{jj}(z_i, z_i) + \mathcal{O}(|h|^{\alpha_\chi}).
\end{aligned} \tag{2.50}$$

Seguidamente, para el tercer término desde 2.47, considerando 2.48 y bajo la hipótesis 2.4

$$\begin{aligned}
\frac{1}{(\delta_n)^{\alpha_\chi}} A_{j,n}^{(2,3)} &= \frac{2^{2H} \sigma^2 C_\chi}{nh^{d+1} (\delta_n)^{\alpha_\chi}} \sum_{l=1}^n K_h^2(z_l - z_i) \sum_{l'=1}^n \|z_l - z_{l'}\|^{\alpha_\chi} \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq 3\delta_n\}} \\
&\leq \frac{2^{2H} \sigma^2 C_\chi (3\delta_n)^{\alpha_\chi}}{nh^{d+1} (\delta_n)^{\alpha_\chi}} \sum_{l=1}^n K_h^2(z_l - z_i) \sum_{l'=1}^n \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq 3\delta_n\}} \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{2^{2H} 3^{\alpha_\chi+d+1} \sigma^2 C_\chi}{h^{d+1}} \int_{\mathbb{R}^{d+1}} K_h^2(z - z_i) dz \\
&= 2^{2H} 3^{\alpha_\chi+d+1} \sigma^2 C_\chi \|K\|_2^2.
\end{aligned} \tag{2.51}$$

Por último, para el cuarto término de 2.47 usamos 2.48 en suma con las hipótesis 2.4 y 2.6

$$\begin{aligned}
\frac{1}{(\delta_n)^{\alpha_\chi + \alpha_k}} A_{j,n}^{(2,4)} &= \frac{2^{2H} \sigma^2 C_K C_\chi}{nh^{d+1} (\delta_n)^{\alpha_\chi + \alpha_k}} \sum_{l=1}^n K_h(z_l - z_i) \sum_{l'=1}^n \|z_l - z_{l'}\|^{\alpha_\chi + \alpha_k} \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq 3\delta_n\}} \\
&\leq \frac{2^{2H} \sigma^2 C_K C_\chi (3\delta_n)^{\alpha_\chi + \alpha_k}}{nh^{d+1} (\delta_n)^{\alpha_\chi + \alpha_k}} \sum_{l=1}^n K_h(z_l - z_i) \sum_{l'=1}^n \mathbb{I}_{\{0 < \|z_l - z_{l'}\| \leq 3\delta_n\}} \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{2^{2H} 3^{\alpha_\chi + \alpha_k + d + 1} \sigma^2 C_K C_\chi}{h^{d+1}} \int_{\mathbb{R}^{d+1}} K_h(z - z_i) dz \\
&= 2^{2H} 3^{\alpha_\chi + \alpha_k + d + 1} \sigma^2 C_K C_\chi.
\end{aligned} \tag{2.52}$$

Así, desde (2.49), (2.50), (2.51), y (2.52) tenemos

$$\begin{aligned} & \frac{1}{nh^{d+1}(\delta_n)^{2H+d+\alpha}} A_{j,n}^{(2)}(z_i) \\ & \stackrel{\leq}{\underset{n \rightarrow \infty, h \rightarrow 0}{\leq}} 2^{2H} 3^{d+1} \sigma^2 \chi_{jj}(z_i, z_i) \|K\|_2^2 + \mathcal{O}(|h|^{\alpha_\chi} \vee (\delta_n)^{\alpha_K} \vee (\delta_n)^{\alpha_\chi} \vee (\delta_n)^{\alpha_\chi + \alpha_K}) \\ & = C_2(H) \chi_{jj}(z_i, z_i) + \mathcal{O}(|h|^{\alpha_\chi} \vee (\delta_n)^{\alpha_K} \vee (\delta_n)^{\alpha_\chi}), \end{aligned} \quad (2.53)$$

donde $C_2(H) = 2^{2H} 3^{d+1} \sigma^2 \|K\|_2^2$.

Finalmente consideramos el caso $\|z_l - z_{l'}\| > 3\delta_n$, dividimos $A_{j,n}^{(3)}(z_i)$ en tres términos:

$$\begin{aligned} & \frac{1}{n^2 h^{2(d+1)} (\delta_n)^{2H+d+\alpha}} A_{j,n}^{(3)}(z_i) \\ & = \frac{1}{n^2 h^{2(d+1)} (\delta_n)^{2H+d+\alpha}} \sum_{1 \leq l \neq l' \leq n} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \mathbb{E}(\varepsilon_l \varepsilon_{l'}) \mathbb{I}_{\{|t_l - t_{l'}| \leq 3\delta_n, \|u_l - u_{l'}\| > 3\delta_n\}} \\ & + \frac{1}{n^2 h^{2(d+1)} (\delta_n)^{2H+d+\alpha}} \sum_{1 \leq l \neq l' \leq n} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \mathbb{E}(\varepsilon_l \varepsilon_{l'}) \mathbb{I}_{\{|t_l - t_{l'}| > 3\delta_n, \|u_l - u_{l'}\| \leq 3\delta_n\}} \\ & + \frac{1}{n^2 h^{2(d+1)} (\delta_n)^{2H+d+\alpha}} \sum_{1 \leq l \neq l' \leq n} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \mathbb{E}(\varepsilon_l \varepsilon_{l'}) \mathbb{I}_{\{|t_l - t_{l'}| > 3\delta_n, \|u_l - u_{l'}\| > 3\delta_n\}} \\ & = A_{j,n}^{(3,1)} + A_{j,n}^{(3,2)} + A_{j,n}^{(3,3)}. \end{aligned} \quad (2.54)$$

En el caso $|t_l - t_{l'}| \leq 3\delta_n$ y $\|u_l - u_{l'}\| > 3\delta_n$ la covarianza $\mathbb{E}(\varepsilon_l \varepsilon_{l'})$ está acotada por

$$\begin{aligned} \mathbb{E}(\varepsilon_l \varepsilon_{l'}) & = \frac{1}{2} (|t_l - t_{l'} + 2\delta_n|^{2H} + |t_l - t_{l'} - 2\delta_n|^{2H} - 2|t_l - t_{l'}|^{2H}) \int_{V(u_l)} \int_{V(u_{l'})} \gamma_{\alpha,d} \|u - u'\|^{-d+\alpha} du du' \\ & \leq (2\delta_n)^{2H} \gamma_{\alpha,d} (\delta_n)^{-d+\alpha} \lambda(V(u_l)) \lambda(V(u_{l'})) \\ & \leq 2^{2H} \gamma_{\alpha,d} \lambda^2(\mathcal{I}^{d-1}) (\delta_n)^{2H+d+\alpha}. \end{aligned} \quad (2.55)$$

Entonces, a partir de (2.54) y (2.55) tenemos

$$\begin{aligned}
A_{j,n}^{(3,1)} &= \frac{1}{n^2 h^{2(d+1)} (\delta_n)^{2H+d+\alpha}} \sum_{1 \leq l \neq l' \leq n} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \mathbb{E}(\varepsilon_l \varepsilon_{l'}) \mathbb{I}_{\{|t_l - t_{l'}| \leq 3\delta_n, \|u_l - u_{l'}\| > 3\delta_n\}} \\
&\leq \frac{2^{2H} \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1})}{n^2 h^{2(d+1)}} \sum_{1 \leq l \neq l' \leq n} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{2^{2H} \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1})}{h^{2(d+1)}} \int_{\mathbb{R}^{d+1}} \int_{\mathbb{R}^{d+1}} \chi_{jj}(z, z') K_h(z - z_i) K_h(z' - z_i) dz dz' \\
&= 2^{2H} \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1}) \int_{\mathbb{R}^{d+1}} \int_{\mathbb{R}^{d+1}} \chi_{jj}(z_i + hz, z_i + hz') K(z) K(z') dz dz' \\
&\stackrel{h \rightarrow 0}{\simeq} 2^{2H} \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1}) \chi_{jj}(z_i, z_i) + 2^{2H} \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1}) |h|^{\alpha_\chi} \int_{\mathbb{R}^{d+1}} \int_{\mathbb{R}^{d+1}} (\|z\| + \|z'\|)^{\alpha_\chi} K(z) K(z') dz dz' \\
&= 2^{2H} \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1}) \chi_{jj}(z_i, z_i) + \mathcal{O}(|h|^{\alpha_\chi})
\end{aligned} \tag{2.56}$$

Ahora, estudiamos el caso $|t_l - t_{l'}| > 3\delta_n$ y $\|u_l - u_{l'}\| \leq 3\delta_n$. Desde (2.18) y (2.19) podemos vincular la covarianza $\mathbb{E}(\varepsilon_l \varepsilon_{l'})$ como sigue

$$\begin{aligned}
\mathbb{E}(\varepsilon_l \varepsilon_{l'}) &= \frac{1}{2} (|t_l - t_{l'} + 2\delta_n|^{2H} + |t_l - t_{l'} - 2\delta_n|^{2H} - 2|t_l - t_{l'}|^{2H}) (\delta_n)^{d+\alpha} \text{Cov} \left(\mathbb{I}_{\{\|u - u_l/\delta_n\| \leq 1\}}, \mathbb{I}_{\{\|u - u_{l'}/\delta_n\| \leq 1\}} \right) \\
&= \frac{1}{2} \left(\int_{t_l - \delta_n}^{t_l + \delta_n} \int_{t_{l'} - \delta_n}^{t_{l'} + \delta_n} 2H(2H - 1) |t - t'|^{2H-2} dt' dt \right) (\delta_n)^{d+\alpha} \text{Cov} \left(\mathbb{I}_{\{\|u - u_l/\delta_n\| \leq 1\}}, \mathbb{I}_{\{\|u - u_{l'}/\delta_n\| \leq 1\}} \right) \\
&\leq H(2H - 1) (\delta_n)^{2H-2} \left(\int_{t_l - \delta_n}^{t_l + \delta_n} \int_{t_{l'} - \delta_n}^{t_{l'} + \delta_n} dt' dt \right) (\delta_n)^{d+\alpha} \text{Var}^{1/2}(\mathbb{I}_{\{\|u - u_l/\delta_n\| \leq 1\}}) \text{Var}^{1/2}(\mathbb{I}_{\{\|u - u_{l'}/\delta_n\| \leq 1\}}) \\
&\leq 4H(2H - 1) \sigma^2 (\delta_n)^{2H+d+\alpha}.
\end{aligned} \tag{2.57}$$

Entonces, a partir de (2.54) y (2.57) tenemos que

$$\begin{aligned}
A_{j,n}^{(3,2)} &= \frac{1}{n^2 h^{2(d+1)} (\delta_n)^{2H+d+\alpha}} \sum_{1 \leq l \neq l' \leq n} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \mathbb{E}(\varepsilon_l \varepsilon_{l'}) \mathbb{I}_{\{|t_l - t_{l'}| > 3\delta_n, \|u_l - u_{l'}\| \leq 3\delta_n\}} \\
&\leq \frac{4H(2H - 1) \sigma^2}{n^2 h^{2(d+1)}} \sum_{1 \leq l \neq l' \leq n} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{2\sigma^2 2H(2H - 1)}{h^{2(d+1)}} \int_{\mathbb{R}^{d+1}} \int_{\mathbb{R}^{d+1}} \chi_{jj}(z, z') K_h(z - z_i) K_h(z' - z_i) dz dz' \\
&\stackrel{h \rightarrow 0}{\simeq} 4H(2H - 1) \sigma^2 \chi_{jj}(z_i, z_i) + \mathcal{O}(|h|^{\alpha_\chi}).
\end{aligned} \tag{2.58}$$

Para el caso $|t_l - t_{l'}| > 3\delta_n$ y $\|u_l - u_{l'}\| > 3\delta_n$, procedemos de manera análoga a los casos anteriores

$$\begin{aligned}
\mathbb{E}(\varepsilon_l \varepsilon_{l'}) &= \frac{1}{2} (|t_l - t_{l'} + 2\delta_n|^{2H} + |t_l - t_{l'} - 2\delta_n|^{2H} - 2|t_l - t_{l'}|^{2H}) \left(\int_{V(u_l)} \int_{V(u_{l'})} \gamma_{\alpha,d} \|u - u'\|^{-d+\alpha} du du' \right) \\
&= \frac{1}{2} \left(\int_{t_l - \delta_n}^{t_l + \delta_n} \int_{t_{l'} - \delta_n}^{t_{l'} + \delta_n} 2H(2H - 1) |t - t'|^{2H-2} dt' dt \right) \left(\int_{V(u_l)} \int_{V(u_{l'})} \gamma_{\alpha,d} \|u - u'\|^{-d+\alpha} du du' \right) \\
&= 4H(2H - 1) (\delta_n)^{2H} \gamma_{\alpha,d} (\delta_n)^{d+\alpha} \lambda^2(\mathcal{S}^{d-1}) \\
&\leq 4H(2H - 1) \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1}) (\delta_n)^{2H+d+\alpha}.
\end{aligned} \tag{2.59}$$

Así, a partir de (2.54) y (2.59)

$$\begin{aligned}
A_{j,n}^{(3,3)} &= \frac{1}{n^2 h^{2(d+1)} (\delta_n)^{2H+d+\alpha}} \sum_{1 \leq l \neq l' \leq n} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \mathbb{E}(\varepsilon_l \varepsilon_{l'}) \mathbb{I}_{\{|t_l - t_{l'}| > 3\delta_n, \|u_l - u_{l'}\| > 3\delta_n\}} \\
&\leq \frac{4H(2H - 1) \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1})}{n^2 h^{2(d+1)}} \sum_{1 \leq l \neq l' \leq n} \chi_{jj}(z_l, z_{l'}) K_h(z_l - z_i) K_h(z_{l'} - z_i) \\
&\stackrel{n \rightarrow \infty}{\simeq} \frac{4H(2H - 1) \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1})}{h^{2(d+1)}} \int_{\mathbb{R}^{d+1}} \int_{\mathbb{R}^{d+1}} \chi_{jj}(z, z') K_h(z - z_i) K_h(z' - z_i) dz dz' \\
&\stackrel{h \rightarrow 0}{\simeq} 4H(2H - 1) \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1}) \chi_{jj}(z_i, z_i) + \mathcal{O}(|h|^{\alpha_\chi}).
\end{aligned} \tag{2.60}$$

Entonces, desde (2.54), (2.56), (2.58) y (2.60) tenemos

$$\begin{aligned}
&\frac{1}{n^2 h^{2(d+1)} (\delta_n)^{2H+d+\alpha}} A_{j,n}^{(3)}(z_i) \\
&\stackrel{n \rightarrow \infty, h \rightarrow 0}{\leq} \left(2^{2H} \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1}) + 4H(2H - 1) \sigma^2 + 4H(2H - 1) \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1}) \right) \chi_{jj}(z_i, z_i) + \mathcal{O}(|h|^{\alpha_\chi}) \\
&= C_3(H) + \mathcal{O}(|h|^{\alpha_\chi}),
\end{aligned} \tag{2.61}$$

donde $C_3(H) = 2^{2H} \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1}) + 4H(2H - 1) \sigma^2 + 4H(2H - 1) \gamma_{\alpha,d} \lambda^2(\mathcal{S}^{d-1})$. Sustituyendo (2.43), (2.53) y (2.61) en la ecuación (2.42), y usando que $2\lambda(\mathcal{S}^{d-1})(\delta_n)^{d+1} = 1/n$ obtenemos

$$\frac{1}{n^2 h^{2(d+1)}} \mathbb{E} \left((X^T \mathcal{W}(z_i) \varepsilon)_j^2 \right) \leq \frac{(C_1(H) + C_2(H)) (\delta_n)^{\frac{2H+d+\alpha}{d+1}}}{n h^{d+1}} + C_3(H) (\delta_n)^{\frac{2H+d+\alpha}{d+1}} \simeq \frac{C(H)}{n^{1+\theta}},$$

Donde $\theta = \frac{2H+\alpha-1}{d+1} > 0$ si $2H + \alpha - 1 > 0$. Así, la convergencia en L^2 , y más aún en probabilidad, se asegura para $2H + d + \alpha > 0$. Para $2H + \alpha - 1 > 0$, el L^2 radio de $\frac{1}{nh^{d+1}} (X^T \mathcal{W}(z_i) \boldsymbol{\varepsilon})_j$ es más rápido que $1/n$; por ejemplo, cuando $H > \frac{1}{2}$ y $\alpha \geq 0$. Una aplicación directa del lema de Borell-Cantelli nos permitió obtener

$$\frac{1}{nh^{d+1}} (X^T \mathcal{W}(z_i) \boldsymbol{\varepsilon})_j \xrightarrow[n \rightarrow \infty]{a.s.} 0.$$

Por teorema de Slutsky y lema 2.3, la convergencia de $|\hat{\beta}(z_i) - \beta(z_i)| \rightarrow 0$, es:

- En probabilidad, para $2H + d + \alpha > 0$ y $d + 1 + \beta > 0$. En particular para $H > 0$, $\alpha > -d$, y $\beta > -(d + 1)$.
- Casi segura, para $2H + \alpha - 1 > 0$ y $\beta > 0$. En particular para $H > 1/2$, $\alpha \geq 0$ y $\beta > 0$.

□

2.7. Simulaciones

En esta sección, presentamos algunos resultados obtenidos durante el proceso de construcción del presente apartado. En este sentido nos referimos a las simulaciones realizadas para modelar el ruido generado por la covarianza 1, las simulaciones de los posibles escenarios que se estimaran mediante nuestro modelo propuesto en la sección 2.1, las estimaciones mediante los diferentes modelos que proponemos y la representación gráfica para la estabilidad, que presenta el Error Cuadrático Medio (ECM) del estimador de mínimos cuadrados ponderado espacio-temporal, propuesto en cada caso planteado para la superficie escogida y que se quiere representar. Cabe destacar, que estas superficies también funcionan como los coeficiente que acompaña a las covariables dentro del modelo espacio-temporal propuesto en cada caso y según las condiciones de regularidad del índice de Hurst escogido espacialmente y temporalmente.

Se simularon cuatro modelos distintos, es decir:

$$Y = \beta_1 * X + \varepsilon_{H_{s4}, H_{t65}}, \quad (2.62)$$

$$Y = \beta_2 * X + \varepsilon_{H_{s4}, H_{t65}}, \quad (2.63)$$

$$Y = \beta_1 * X + \varepsilon_{H_{s4}, H_{t90}}, \quad (2.64)$$

y

$$Y = \beta_2 * X + \varepsilon_{H_{s4}, H_{t90}}, \quad (2.65)$$

de donde Y es la variable dependiente o respuesta, según corresponda en cada modelo mencionado anteriormente, $\beta_1 = 1 + (4(x+y)/12)$ un plano con cierto gradiente de inclinación, $\beta_2 = 1 + (36 - (6 - (25x)/2)^2)(36 - (6 - (25y)/2)^2)/(324 * 8)$ una superficie curva, ambos escogidas basándonos en ver Que et al. (2020), X es una matriz de covariables simulada a través de un proceso espacio-temporal autoregresivo denotado por $(STAR(1,1))$, al cual nos referiremos de esta manera en lo que sigue, es decir $X = STAR(1,1)$. Además, ε_{H_s, H_t} representa el ruido fraccionario coloreado con índice Hurst espacial de 0,40 e índice de Hurst temporal de 0,65 y de igual forma para ocurre para ε_{H_s, H_t} .

Es importante resaltar que β_1 y β_2 , además de ser los coeficientes que acompañan a las superficies simuladas, representan también la superficie que estima cada modelo mencionado.

Por otra parte, se verifico la pertinencia de β_1 y β_2 dentro de los modelos simulados, al igual que la bondad del ajuste del modelo simulado en contraste con el modelo estimado según el coeficiente escogido dentro en cada modelo. Esto último, mediante un análisis de los residuales que nos conllevo a representar la estabilidad del ECM para β_1 y β_2 mediante iteraciones.

Partiremos la descripción de los resultados mencionados, mostrando los gráficos de la covarianza del ruido espacial, considerando $H_s = 0,40$ y $H_t = 0,65$ como el índice de Hurst espacial y temporal respectivamente, y se presentan en tres tiempo distintos $t = 1$, $t = 50$ y $t = 100$. Seguidamente, las simulaciones de los modelos considerando β_1 y β_2 con sus respectivos análisis residuales, las estimaciones realizadas para modelo simulado y la representación gráfica para la estabilidad del ECM de los modelos con los índices mencionado. Así, bajo el mismo orden que lo anterior, se realizó el análisis de los dos modelos faltantes, pero con $H_s = 0,40$ y $H_t = 0,90$ como índice espacial y temporal respectivamente.

2.7.1. Ruido coloreado con ($H_s = 0,40$ y $H_t = 0,65$)

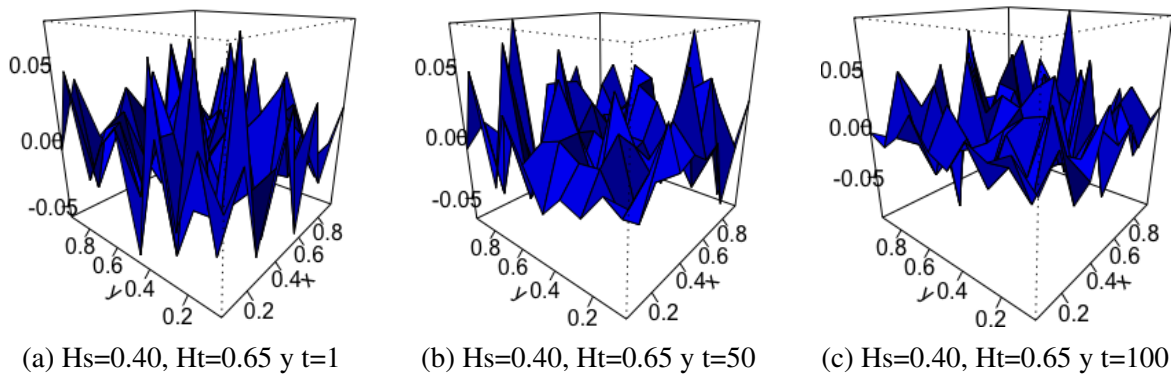


Figura 2.1 Ruido coloreado espacial

En la figura 2.1 podemos observar la simulación del ruido coloreado espacial para los instantes de tiempo $t = 1$, $t = 50$ y $t = 100$, considerando una constante de Hurst espacial $H_s = 0,40$ y temporal $H_t = 65$. Cada uno de los gráfico muestra una superficie con mucha rugosidad indistintamente del tiempo, ésto destaca la similitud que presenta el ruido coloreado espacial durante los primeros cien segundos. Además cada simulación claramente oscila alrededor del plano constantemente igual a cero, lo cual corrobora la hipótesis 2.1.

2.7.2. Simulación: modelo ($Y = \beta_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$)

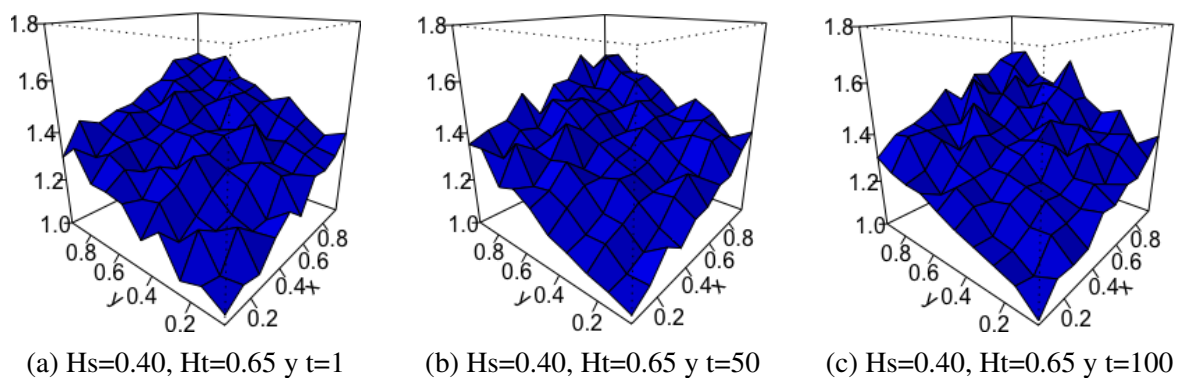


Figura 2.2 Modelo $Y = \beta_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$

La figura 2.2 muestra como la superficie plana β_1 es perturbada por un ruido coloreado en espacio con índice de Hurst $H_s = 0,40$ y fraccionaria en tiempo con índice $H_t = 0,65$. Esto, fue considerado para tres instantes de tiempo distintos $t = 1$, $t = 50$ y $t = 100$ respectivamente. Note que, las imágenes correspondientes a los distintos tiempos para las gráficas de las superficies son muy similares, es decir, mantienen un comportamiento similar para la rugosidad del terreno, el cual es un posible escenario para realizar estimaciones con el modelo que se plantea en éste caso. Más aún, los gráficos presentan baja variabilidad entre los tiempos escogidos para ser representados.

En los cuadros que se presentaran a continuación, mostraremos la pertinencia de incorporar el coeficiente escogido en el modelo que simulamos.

Results of Global Regression				
Residuals:				
Mín	1Q	Median	3Q	Máx
-0,33	-0,10	-0,00129	0,10	0,39
Coefficients:				
Estimate	Std. Error	t value	$Pr(> t)$	
1,33330	0,00139	959,1	$< 2e^{-16}$	
Residual standard error: 0,139 on 9999 degrees of freedom				
Multiple R-squared: 0,9892				
Adjusted R-squared: 0,9892				
F-statistic: $9,199e^5$ on 1 and 9999 DF, p-value: $< 2,2e^{-16}$				
Extra Diagnostic information				
Residual sum of squares: 193,276				
Sigma(hat): 0,1390377				
AIC: -11079,44				
AICc: -11079,44				
BIC: -21046,6				

Cuadro 2.1 Resultados de la regresión global $Y = \beta_1 * STAR(1, 1) + \varepsilon_{H_{34}, H_{165}}$

Desde el cuadro 2.1, se muestra el resultado del análisis de regresión global para modelo propuesto. Allí, β_1 se considera constante para cada instante de tiempo. La significancia estadística del coeficiente dentro del modelo se hace notar pues para $\alpha = 0,001$, $\alpha = 0,01$, $\alpha = 0,05$, $\alpha = 0,1$ y $\alpha = 1$ el valor crítico igual $2e^{-16}$ es menor siempre. Es decir, bajo cualquiera de estos α considerados significativo incluir β_1 como coeficiente regresor. Más aún, a partir del R^2 múltiple cuyo valor es 0,9892, podemos afirmar que predice un poco más del 98 % de los datos a pesar de la perturbación que se suma con el ruido fraccionario coloreado. Finalmente, la bondad del ajuste del modelo en concordancia con el nivel de complejidad, evidencia que para el valor $AIC = -11079,44$ y $BIC = -21046,6$ el modelo es bastante adecuado. Sin embargo, el próximo cuadro nos muestra el análisis de regresión para el modelo ponderado y es probable que mejore los resultados de la regresión global.

Results of Geographically Weighted Regression				
Model calibration information				
Kernel function: Gaussian				
Adaptive bandwidth: 232 (number of nearest neighbours.)				
Regression points: the same locations as observations are used.				
Distance metric: Euclidean distance metric is used.				
Summary of GWR coefficient estimates:				
Mín	1Q	Median	3Q	Máx
1,07	1,24	1,33	1,43	1,60
Diagnostic information				
Number of data points: 10000				
Effective number of parameters (2trace(S) - trace(S'S)): 28,03505				
Effective degrees of freedom (n-2trace(S) + trace(S'S)): 9971,965				
AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): -39451,39				
AIC (GWR book, Fotheringham, et al. 2002, GWR p. 96, eq. 4.22): -39472,71				
BIC (GWR book, Fotheringham, et al. 2002, GWR p. 61, eq. 2.34): -49314,79				
Residual sum of squares: 11,28268				
R-square value: 0,9421051				
Adjusted R-square value: 0,9419423				

Cuadro 2.2 Resultados de la regresión global $Y = \beta_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$

Ahora bien, el presente cuadro 2.2 muestra el análisis de la regresión ponderada de acuerdo al modelo considerado, es decir, en este caso β_1 para cada instante de tiempo t es estimado y no se deja fijo como en la regresión global. Así, se pudo apreciar una mejora bastante grande en cuanto a la bondad de ajuste del modelo y la complejidad del mismo, pues para este caso y en comparación al modelo global el $AIC = -39472,71$ al igual que el $BIC = -49314,79$ son más pequeños aún. Esto, indica la buena adecuación del modelo simulado. Más aún, la suma de los residuales en la regresión ponderada es notoriamente más baja que en la regresión global y aunque el R^2 disminuye un poco en cuanto a el porcentaje de variabilidad que capta, por los motivos previos, se concluye que la regresión ponderada estima mejor que la regresión global. Note que, el cuadro 2.2 nos indica el kernel ocupado para la estimación, junto con el ancho de banda adecuado a la cantidad de vecinos más cercanos óptima.

2.7.3. Estimación: modelo ($\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$)

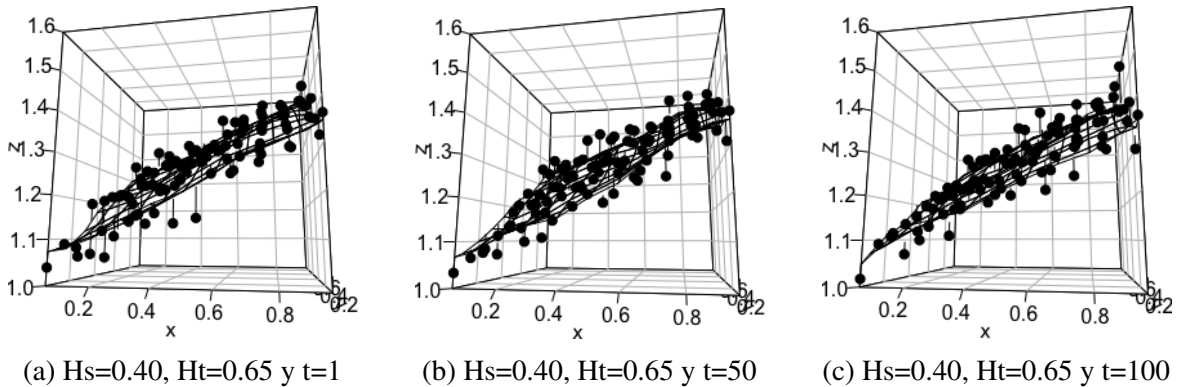


Figura 2.3 $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$

A partir de las superficies simuladas anteriormente, las cuales sirven como referencia de valores reales de superficies que podrían ser estimadas. Proponemos $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$ como el estimador de las observaciones en la figura 2.3. Así, los puntos de color negro ubicados sobre el plano levemente perturbado con cierto gradiente de inclinación, son las observaciones estimadas en distintos instante de tiempo considerados sobre dicha superficie. Note que, para los tres casos (a), (b) y (c), las estimaciones son bastante similares entre si y muy cercanas a la superficie simulada. Esto, corrobora lo mencionado en el análisis del modelo ponderado en relación a su buena bondad del ajuste. Sin embargo, una mirada al gráfico de los residuos que se presentan a continuación y que son resultado del modelo ajustado, nos ayudara en la verificación de que estos no presentan tanta variabilidad respecto a la superficie que estiman y se comportan aproximadamente normal.

2.7.4. Residuos: modelo $(\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}})$

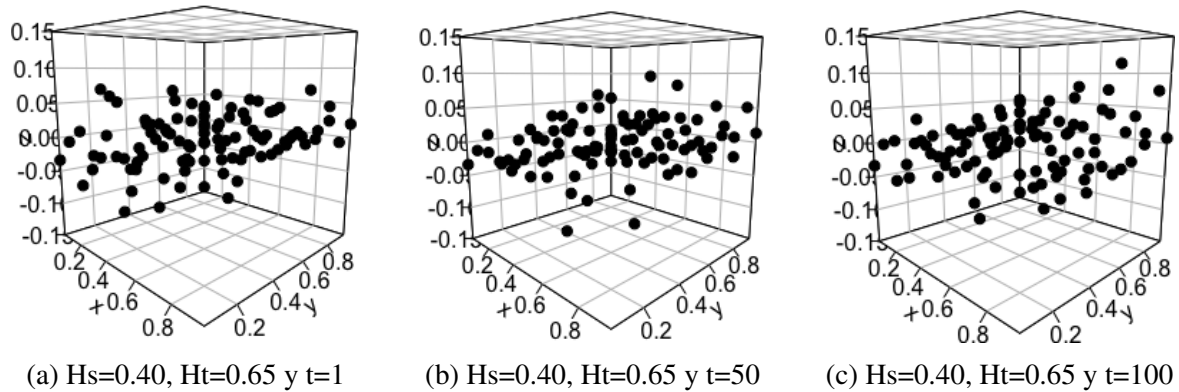


Figura 2.4 Residuos: modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$

Como se puede apreciar en los gráficos (a),(b) y (c) de la figura 2.4, los residuos obtenidos entre el modelo real (modelo simulado) y el modelo estimado presentan poca variabilidad para los tres instantes de tiempo que fueron fijados. Además, éstas variaciones residuales ocurren alrededor del plano constantemente igual a cero, por lo que podríamos casi afirmar que los residuos tienen un comportamiento aproximadamente normal. Por otra parte, el siguiente gráfico de las densidades para los residuos nos corroboran parte de la información mencionada.

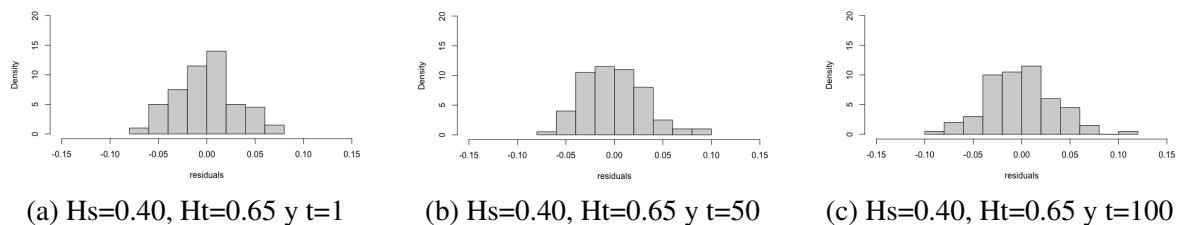


Figura 2.5 Histogramas de los residuos para el modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$

Tal como se dijo anteriormente, las densidades de los residuos presentan un comportamiento aproximadamente normal, muestra de esto se puede observar en la gráfico de barra de la figura 2.5, el cual muestra los histogramas de los residuos para los tres tiempo considerados $t = 1, t = 50$ y $t = 100$. Cabe destacar que si bien muestran formas acampanadas, podrían presentar problemas en las colas debido a su decaimiento pesado en los extremos. El gráfico cuantil-cuantil para cada instante de tiempo considerado durante el desarrollo, nos podría esclarecer la sospecha.

Desde la figura 2.6 se puede corroborar el hecho mencionado al analizar el histograma para los residuales en los distintos instantes de tiempo, pues si bien presentan una distribución

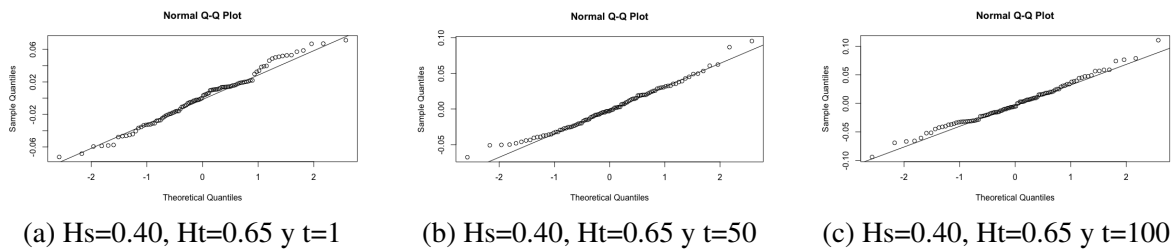


Figura 2.6 Gráficos de QQ-plot: modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$

aproximadamente normal, también muestran un decaimiento lento de las colas. Esto, se puede evidenciar en cada gráfico cuantil-cuantil, donde la mayoría de los puntos se pega a la línea diagonal, la cual compara los cuantiles teóricos de la distribución normal con los cuantiles de las distribuciones de los residuales para el modelo usado y así, notamos como algunos de los puntos a los extremos se desvían de dicha diagonal. Sin embargo, para los fines del estudio presentado, no implica mayor relevancia este hecho, pues el error cuadrático medio entre el estimador planteado y los valores reales es bastante estable. Muestra de esto, se presenta en el próximo gráfico.

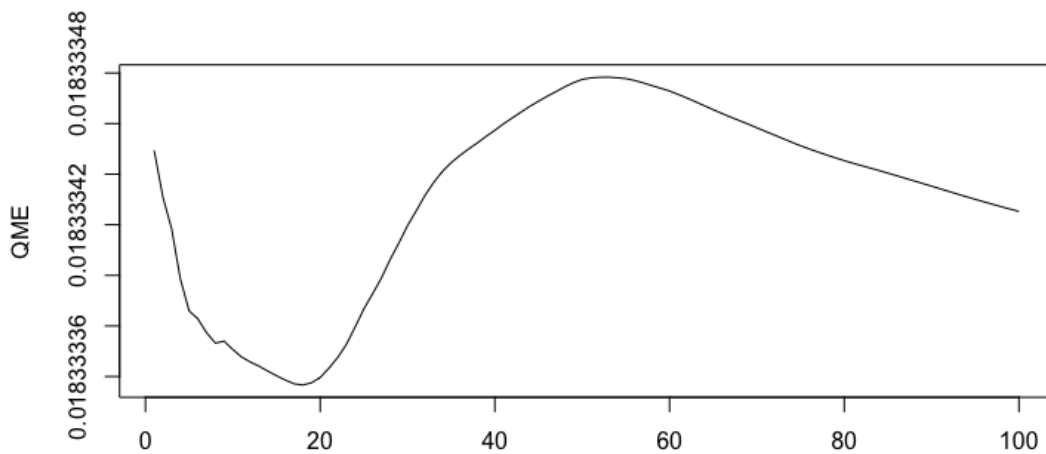


Figura 2.7 (ECM): modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$

La figura 2.7 muestra la gráfica para el comportamiento para el ECM asociado al estimador $\hat{\beta}_1$, este error fue calculado por iteraciones donde se iban almacenando los t , es decir, en la primera iteración se estimó el parámetro con 100 observaciones generando la primera grilla o $t = 1$. En la segunda iteración se consideraron 200 puntos ($t = 1$ y $t = 2$) y así hasta completar las 10000 observaciones que representan una grilla ($10 \times 10 \times 100$). Por otra parte, se puede notar en la gráfica, que los valores resultantes al calcular el ECM en cada tiempo $t = 1, \dots, t = 100$ son bastante pequeños y varían notablemente para los primeros valores, pero para valores más alejado del cero se evidencia un comportamiento estable a lo largo

del tiempo, tal como plantea desde 2.1. Finalmente, esta observación sobre la estabilidad del ECM para estimador $\hat{\beta}_1$, se puede reforzar observando los gráficos de caja asociados al comportamiento que presenta el ECM a lo largo de las 10000 observaciones.

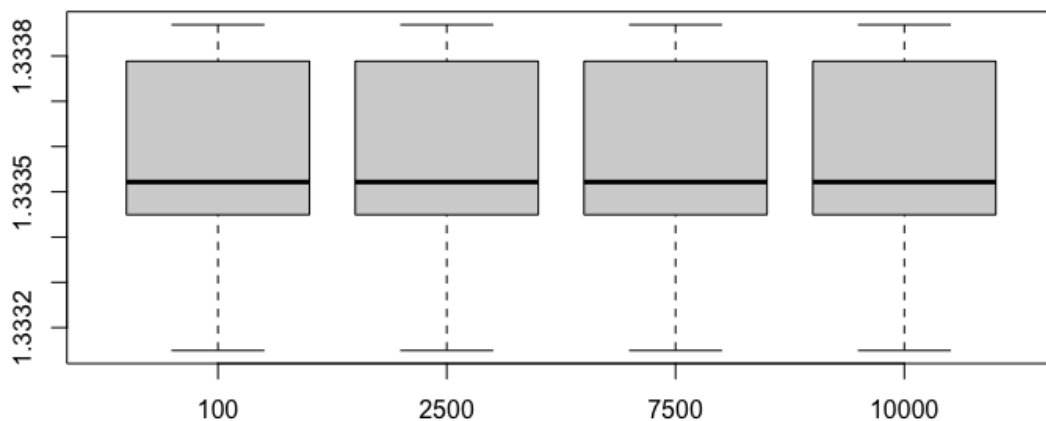


Figura 2.8 Box-plot para $\hat{\beta}_1$

Finalmente, observando el gráfico de caja, es fácil corroborar que el comportamiento del ECM para $\hat{\beta}_1$ a medida que se itera para las 10000 observaciones, es bastante estable.

2.7.5. Simulación: modelo ($Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$)

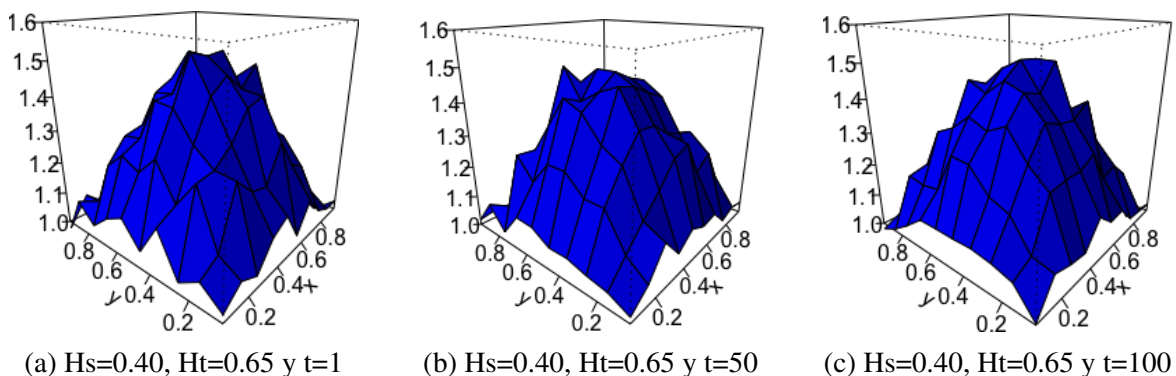


Figura 2.9 Modelo $Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$

La figura 2.9 muestra como la superficie curva escogida y denotada por β_2 es perturbada por un ruido coloreado en espacio con índice de Hurst $H_s = 0,40$ y fraccionaria en tiempo con índice $H_t = 0,65$ para tiempos $t = 1, t = 50$ y $t = 100$ respectivamente. Cabe destacar, que las superficies para los distintos tiempos son muy similares entre si, manteniendo un comportamiento parecido en cuanto a la rugosidad del terreno simulado. Esto indica la baja

variabilidad para los distintos t que se escojan. Veamos ahora, el análisis de regresión global y regresión ponderada respectivamente asociado a este modelo.

Results of Global Regression				
Residuals:				
Mín	1Q	Median	3Q	Máx
-0,32	-0,14	-0,01514	0,13	0,39
Coefficients:				
Estimate	Std. Error	t value	Pr(> t)	
1,20	0,002	763,1	< $2e^{-16}$	
Residual standard error: 0,1579 on 9999 degrees of freedom				
Multiple R-squared: 0,9831				
Adjusted R-squared: 0,9831				
F-statistic: $5,824e^5$ on 1 and 9999 DF, p-value: < $2,2e^{-16}$				
Extra Diagnostic information				
Residual sum of squares: 249,3051				
Sigma(hat): 0,1579098				
AIC: -8533,858				
AICc: -8533,856				
BIC: -18501,02				

Cuadro 2.3 Resultados de la regresión global $Y = \beta_2 * STAR(1, 1) + \varepsilon_{H_{34}, H_{165}}$

Desde el cuadro 2.3 se muestra el resultado para el análisis de regresión global del modelo propuesto. Al igual que en el modelo asociado a β_1 , el coeficiente β_2 en este caso, se considera constante para cada instante de tiempo. La significancia estadística del coeficiente dentro del modelo se hace notar, pues para $\alpha = 0,001$, $\alpha = 0,01$, $\alpha = 0,05$, $\alpha = 0,1$ y $\alpha = 1$ el valor crítico $2e^{-16}$ es siempre menor a cualquiera de estos niveles. Es decir, bajo cualquiera de estos α , es considerado significativo incluir a β_2 como coeficiente regresor. Más aún, a partir del R^2 múltiple cuyo valor es 0,9831, podemos afirmar que el modelo predice alrededor del 98% de la variabilidad en los datos. Esto, aún y cuando se le sume la perturbación del ruido fraccionario coloreado. Por otra parte, la bondad del ajuste del modelo en concordancia con el nivel de complejidad, se evidencian como buenas para el modelo, al observar un valor bajo en $AIC = -8533,858$ y $BIC = -21046,6$. Sin embargo, el próximo cuadro nos muestra el análisis de regresión para el modelo ponderado, el cual es probable que mejore los resultados de la regresión global.

Results of Geographically Weighted Regression				
Model calibration information				
Kernel function: Gaussian				
Adaptive bandwidth: 232 (number of nearest neighbours.)				
Regression points: the same locations as observations are used.				
Distance metric: Euclidean distance metric is used.				
Summary of GWR coefficient estimates:				
Mín	1Q	Median	3Q	Máy
1,02	1,11	1,19	1,30	1,45
Diagnostic information				
Number of data points: 10000				
Effective number of parameters (2trace(S) - trace(S'S)): 28,03761				
Effective degrees of freedom (n-2trace(S) + trace(S'S)): 9971,962				
AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): -28990,24				
AIC (GWR book, Fotheringham, et al. 2002,GWR p. 96, eq. 4.22): -29011,56				
BIC (GWR book, Fotheringham, et al. 2002,GWR p. 61, eq. 2.34): -38853,62				
Residual sum of squares: 32,11694				
R-square value: 0,8699444				
Adjusted R-square value: 0,8695787				

Cuadro 2.4 Resultados de la regresión global $Y = \beta_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$

El cuadro 2.4 muestra el análisis de la regresión ponderada de acuerdo con el modelo $Y = \beta_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$, es decir, en este caso β_2 para cada instante de tiempo t es estimado y no se deja fijo como en la regresión global. Así, es posible apreciar una mejora bastante grande en cuanto a la bondad de ajuste del modelo en concordancia con la complejidad del mismo, pues para este caso y en comparación al modelo global el $AIC = -29011,56$ al igual que el $BCI = -38853,62$ son más pequeños aún. Esto, indica la buena adecuación del modelo simulado. Más aún, la suma de los residuales en la regresión ponderada es notoriamente más baja que en la regresión global y aunque el R^2 ajustado disminuye un poco en cuanto a el porcentaje de variabilidad que capta, por los motivos previos, se concluye que la regresión ponderada estima mejor que la regresión global. Note que el cuadro 2.4 nos indica el kernel ocupado para la estimación, junto con el ancho de banda adecuado a la cantidad de vecinos más cercanos óptima.

2.7.6. Estimación: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$)

Usando las superficies simuladas anteriormente, las cuales sirven como referencia de valores reales a ser estimados. Proponemos $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$ como el estimador de las observaciones en la figura 2.10. Ahora bien, los puntos de color negro ubicados sobre

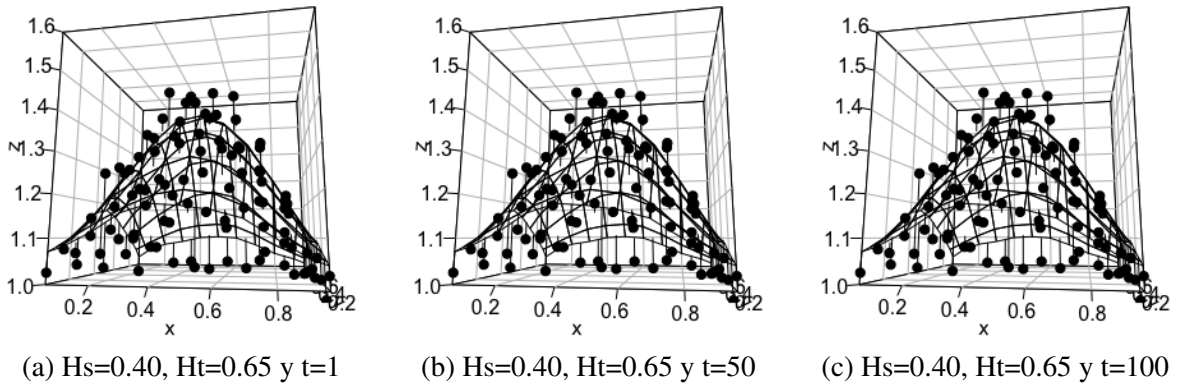


Figura 2.10 \hat{y} para β_2

la superficie curva levemente perturbada, son las observaciones estimadas sobre superficie en distintos instante de tiempo considerados. Luego, para los tres casos (a), (b) y (c), las estimaciones son bastante similares entre si y muy cercanas a la superficie simulada. Este hecho, corrobora que el modelo ponderado en relación a una bondad de ajuste es bastante bueno. Sin embargo, una mirada al gráfico de los residuos para el modelo ajustado, nos ayudara en la verificación que estos no presentan tanta variabilidad respecto a la superficie que estiman y se comportan aproximadamente normal.

2.7.7. Residuos: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$)

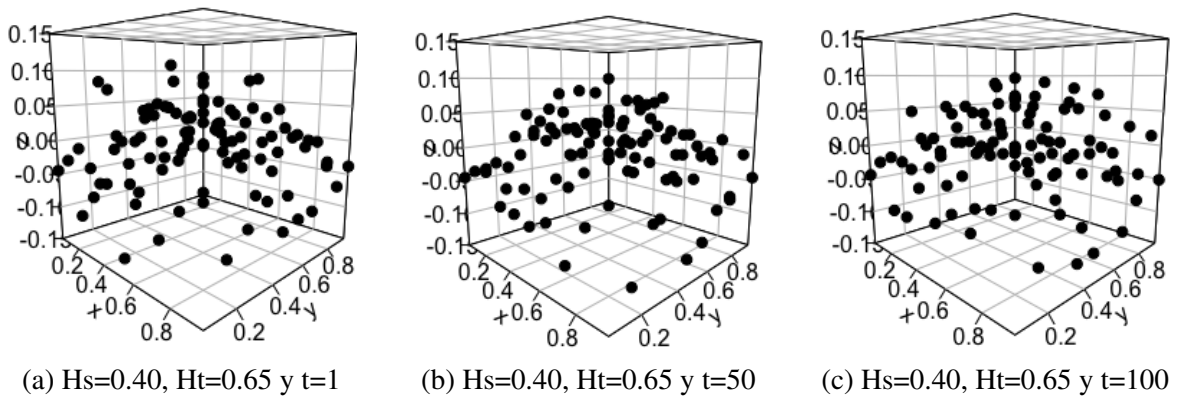
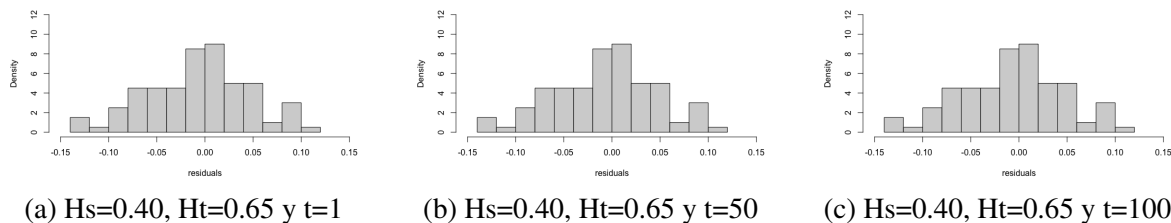


Figura 2.11 Residuos: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t65}}$)

Los gráficos (a),(b) y (c) de la figura 2.11, muestran los residuos obtenidos entre el modelo real (modelo simulado) y el modelo estimado, los cuales presentan mediana variabilidad para los tres instantes de tiempo que fueron fijados, es decir, para $t = 1, t = 50$ y $t = 100$. Cabe destacar, que éstas variaciones residuales ocurren alrededor del plano constantemente

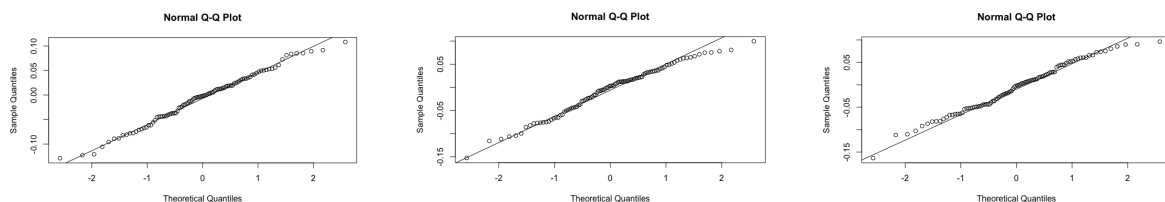
igual a cero, por lo que podríamos casi afirmar que los residuos tienen un comportamiento aproximadamente normal. Sin embargo, basándonos en análisis anteriores podríamos afirmar que presentan problemas de ajuste a dicha distribución en las colas.



(a) $H_s=0.40$, $H_t=0.65$ y $t=1$ (b) $H_s=0.40$, $H_t=0.65$ y $t=50$ (c) $H_s=0.40$, $H_t=0.65$ y $t=100$

Figura 2.12 Histogramas de los residuos para el modelo $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$

La figura 2.12 muestra el histograma para cada una de las distribuciones residuales en los instantes de tiempo $t = 1, t = 2$ y $t = 3$ del modelo $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$. Estas distribuciones, presentan una similitud con la distribución normal, sin embargo se hace más evidente la cola pesada de éstas en comparación al modelo trabajado con β_1 . Podríamos decir, que ello se debe al cambio de superficie o coeficiente dentro modelo propuesto para este ejemplo y que asume las mismas condiciones de regularidad en el índice de Hurst tanto espacial como temporal. Una manera de chequear que en efecto los residuales presentan problemas en las colas es mirando los gráficos cuantil-cuantil entre la distribución normal teórica y la distribución de los residuales.



(a) $H_s=0.40$, $H_t=0.65$ y $t=1$ (b) $H_s=0.40$, $H_t=0.65$ y $t=50$ (c) $H_s=0.40$, $H_t=0.65$ y $t=100$

Figura 2.13 Gráficos de QQ-plot: modelo $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$

Desde la figura 2.13 se puede corroborar el hecho mencionado en el análisis anterior, pues a pesar de que las gráficas de las residuales presentan una distribución aproximadamente normal, también muestran un decaimiento lento de las colas. Así, podemos observar en cada gráfico cuantil-cuantil para los tiempos $t = 1, t = 50$ y $t = 100$ respectivamente, como la mayoría de los puntos en el centro de la diagonal están más cercanos a ésta, pero para los puntos que están más alejados del centro se evidencia una ligera separación con respecto a la diagonal. Esto, deja en evidencia gráfica que las colas de las distribución decaen de forma más pesada, sin embargo, es posible que al aumenta el número de puntos de observación dentro de la simulación los residuos se ajusten con mayor precisión a una distribución normal.

Ahora bien, como ya se ha mencionado antes, el comportamiento evidenciado de los residuos para este caso tampoco afecta los fines de la consistencia para el estimador, pues al mirar el ECM para las iteraciones del estimador $\hat{\beta}_2$ se comprueba la estabilidad de éste.

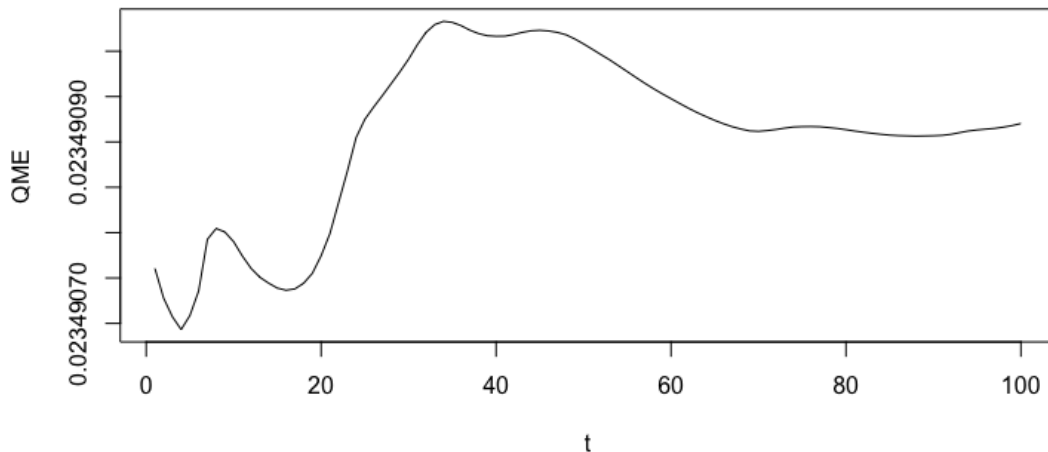


Figura 2.14 (ECM): modelo $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t65}}$

A partir de la figura 2.14 se muestra la gráfica para el comportamiento para el ECM asociado al estimador $\hat{\beta}_2$, este error fue calculado por iteraciones y de forma análoga al explicado en 2.7. Así, al observar la gráfica del ECM para el estimador, particularmente los valores que toma para cada t desde $t = 1, \dots, t = 100$ se aprecia que son bastante pequeños y varían notoriamente para los primeros 60 tiempos aproximadamente, ya que a partir de dicho instante los valores empiezan estabilizar su comportamiento permitiendo evidenciar gráficamente la consistencia del estimador.

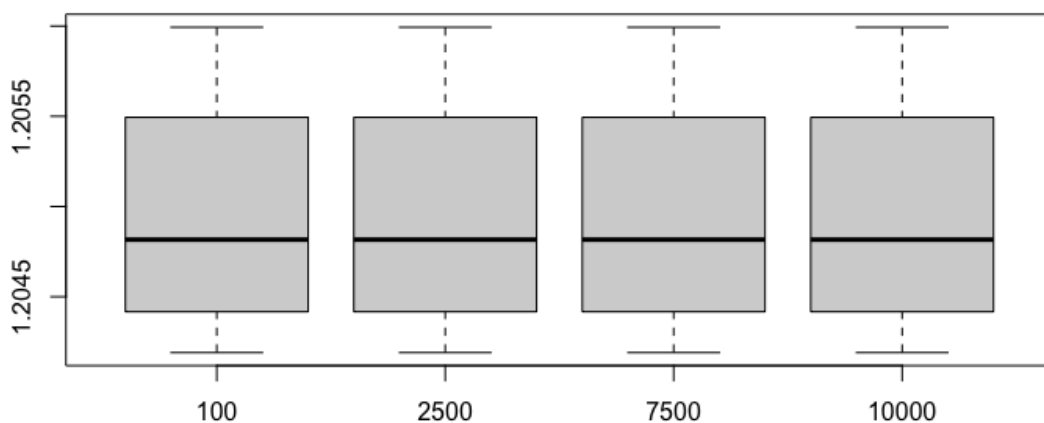


Figura 2.15 Box-plot para $\hat{\beta}_2$

Finalmente y como apoyo a lo comentado en cuanto al comportamiento estable del estimador para el caso particular, la figura 2.15 evidencia como los datos mantienen el rango de valores exactamente igual para los 10000 puntos de tiempo iterados, lo cual asegura un comportamiento de estabilidad para el estimador β_2 considerando un modelo con ruido fraccionario coloreado con índice de Hurst $H_s = 0,40$ espacial y $H_t = 0,65$ temporal.

2.7.8. Ruido coloreado con ($H_s = 0,40$ y $H_t = 0,90$)

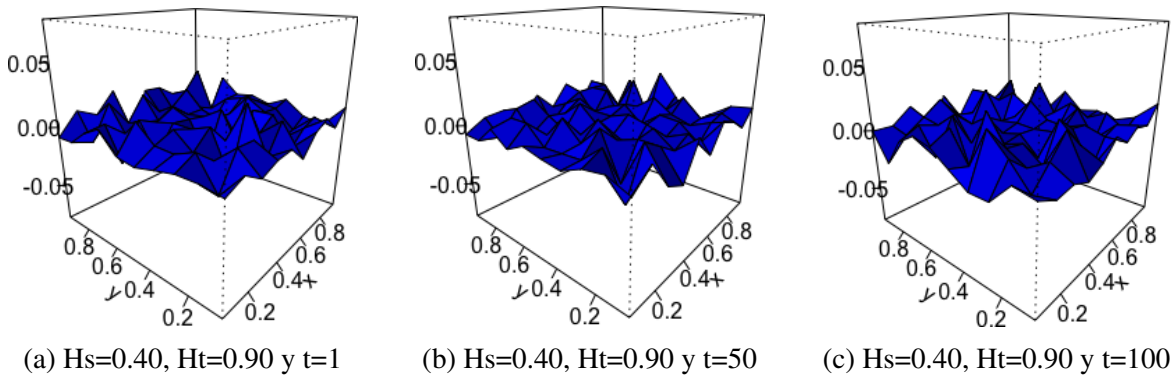


Figura 2.16 Ruido coloreado espacial

En la figura 2.16 podemos observar la simulación del ruido coloreado espacial para los instantes de tiempo $t = 1$, $t = 50$ y $t = 100$ respectivamente. Se consideramos una constante de Hurst espacial $H_s = 0,40$ y temporal $H_t = 0,90$ para éste caso. Cada uno de los gráficos muestra una superficie con rugosidad media en relación al ruido usado anteriormente y cuya diferencia era sólo respecto al índice de Hurst temporal, pues era menos cercano uno. Cabe destacar, que la rugosidad de menor intensidad ocurre cuando los valores del índice Hurst H están más cercanos a uno. Por otra parte, cada simulación claramente oscila alrededor del plano constantemente igual a cero, así comparte un comportamiento similar con el ruido generado para $H_t = 0,65$.

2.7.9. Simulación: modelo ($Y = \beta_1 * STAR(1, 1) + \epsilon_{H_s, H_t}$)

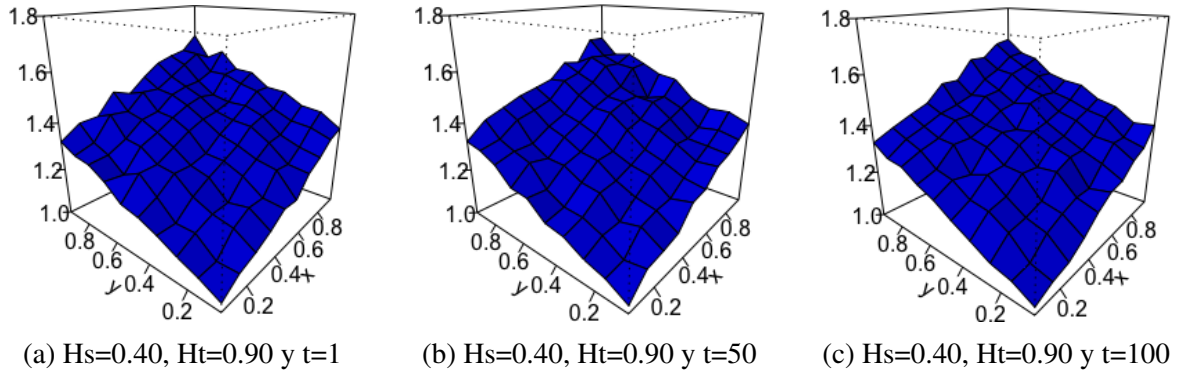


Figura 2.17 Modelo $Y = \beta_1 * STAR(1, 1) + \epsilon_{H_s, H_t}$

En este caso la figura 2.17 muestra como la superficie plana β_1 es perturbada por un ruido coloreado espacialmente con índice de Hurst $H_s = 0,40$ y fraccionario en tiempo con índice $H_t = 0,90$. De manera análoga se consideraron tres instantes de tiempo distintos $t = 1$, $t = 50$ y $t = 100$ respectivamente para realizar las simulaciones. Note que, las imágenes correspondientes a los distintos tiempos para las gráficas de las superficies son muy similares entre si, pero a diferencia de las consideradas con un ruido fraccionario menos cercano a uno, éstas producen un efecto de menor rugosidad en las imágenes, lo que podría significar, en la práctica, que nos encontramos ante un escenario de terreno menos irregular para realizar la estimaciones. Ahora bien, en los cuadros que se presentaran a continuación, evaluaremos la significancia estadística de incorporar el coeficiente β_1 escogido en el modelo de regresión ponderado que simulamos. Sin embargo, cómo el cambio se realizo principalmente en la componente fraccionaria temporal del ruido, algunos comportamientos del modelo de regresión pueden ser similares a los anteriores.

Results of Global Regression				
Residuals:				
Mín	1Q	Median	3Q	Máx
-0,31	-0,10	-0,00216	0,34	0,39
Coefficients:				
Estimate	Std. Error	t value	Pr(> t)	
1,20	0,002	763,1	< $2e^{-16}$	
Residual standard error: 0,1357 on 9999 degrees of freedom				
Multiple R-squared: 0,9898				
Adjusted R-squared: 0,9898				
F-statistic: $9,6684e^5$ on 1 and 9999 DF, p-value: < $2,2e^{-16}$				
Extra Diagnostic information				
Residual sum of squares: 184,1402				
Sigma(hat): 0,1357119				
AIC: -11563,66				
AICc: -11563,66				
BIC: -21530,82				

Cuadro 2.5 Resultados de la regresión global $Y = \beta_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$

Desde el cuadro 2.5 se muestra el resultado para el análisis de regresión global del modelo propuesto en este caso. Cabe destacar, que igual a los modelos globales que se analizaron anteriormente, el coeficiente β_1 se mantiene constante a lo largo de las iteraciones en el tiempo. Por otra parte, la significancia estadística del coeficiente dentro del modelo se hace notar nuevamente, pues para $\alpha = 0,001$, $\alpha = 0,01$, $\alpha = 0,05$, $\alpha = 0,1$ y $\alpha = 1$ el valor crítico $2e^{-16}$ es siempre menor a cualquiera de estos niveles. Es decir, bajo cualquiera de estos α , es considerado significativo incluir a β_1 como coeficiente regresor en el modelo. Más aún, la bondad del ajuste para el modelo al igual que la complejidad de su estructura se mantienen en equilibrio debido a que el $AIC = -11563,66$ y $BIC = -21530,82$ son lo suficientemente pequeños. Por otra parte, el R^2 múltiple es igual a 0,9898, es decir, dicho modelo es capaz de predecir el 98% de la variabilidad de los datos sobre la superficie.

Results of Geographically Weighted Regression				
Model calibration information				
Kernel function: Gaussian				
Adaptive bandwidth: 147 (number of nearest neighbours.)				
Regression points: the same locations as observations are used.				
Distance metric: Euclidean distance metric is used.				
Summary of GWR coefficient estimates:				
Mín	1Q	Median	3Q	Máx
1,07	1,24	1,33	1,43	1,60
Diagnostic information				
Number of data points: 10000				
Effective number of parameters (2trace(S) - trace(S'S)): 28,03761				
Effective degrees of freedom (n-2trace(S) + trace(S'S)): 9971,962				
AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): -55378,38				
AIC (GWR book, Fotheringham, et al. 2002, GWR p. 96, eq. 4.22): -55399,7				
BIC (GWR book, Fotheringham, et al. 2002, GWR p. 61, eq. 2.34): -65241,76				
Residual sum of squares: 2,294626				
R-square value: 0,9877775				
Adjusted R-square value: 0,9877432				

Cuadro 2.6 Resultados de la regresión global $Y = \beta_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$

El cuadro 2.6 muestra el análisis de la regresión ponderada de acuerdo con el modelo $Y = \beta_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$, es decir, en este caso β_1 para cada instante de tiempo t es estimado y no se deja fijo como en la regresión global. Además, en cuanto a la bondad de ajuste del modelo en concordancia y la complejidad necesaria para la estructura de este, los valores obtenidos para $AIC = -55399,7$ y para $BCI = -65241,76$ son más pequeños que los considerados en el modelo global. Esto, indica la buena adecuación del modelo simulado. Más aún, la suma de los residuales al cuadrado en la regresión ponderada es más baja que en la regresión global. Finalmente y como otro aspecto a favor del modelo ponderado, note que el R^2 ajustado disminuye poco en relación al modelo global. Así, las estimaciones de modelo ponderado son más precisas en cuanto a las estimaciones del modelo global.

2.7.10. Estimación: modelo ($\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)

Usando las superficies simuladas anteriormente, las cuales nos sirven como referencia de valores reales sobre superficies que podrían ser estimadas. Se propuso como estimador de $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$ como el estimador de las observaciones en la figura 2.18. De manera que, los puntos de color negro ubicados sobre el plano levemente perturbado con cierto gradiente de inclinación, son las observaciones estimadas en distintos instante de

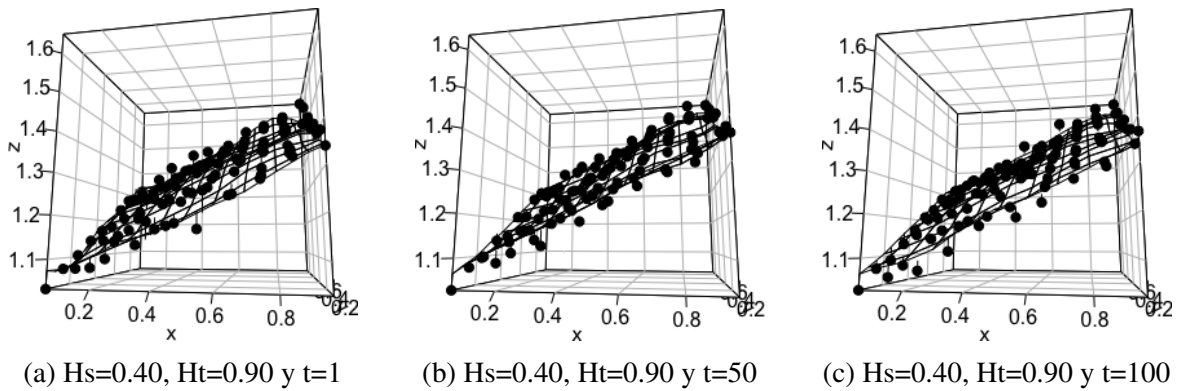


Figura 2.18 Estimación: modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$

tiempo considerados sobre la superficie simulada. Cabe destacar que, para las tres gráficas (a), (b) y (c) presentadas en 2.18, las estimaciones son bastante similares entre sí y muy cercanas a la superficie simulada. Esto, reafirma el hecho de que el modelo ponderado en relación a su buena bondad del ajuste es bastante bueno. Sin embargo, realizaremos una mirada a los gráfico de los residuos que se presentan posteriormente y que son producto del modelo ajustado. Así, podremos revisar con un poco más de detalle que dichos residuos no presentan tanta variabilidad respecto a la superficie que estiman y se comportan aproximadamente normal.

2.7.11. Residuos: modelo ($\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)

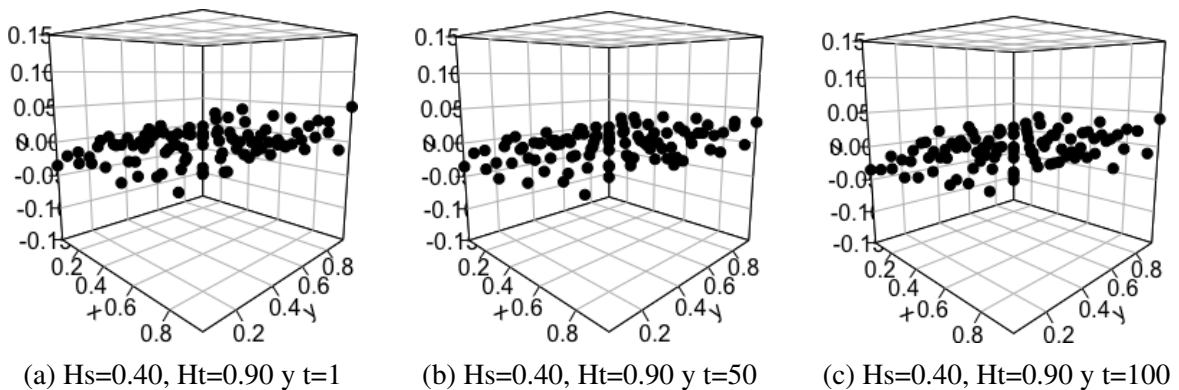


Figura 2.19 Residuos: modelo ($\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)

A partir de los gráficos presentados en la figura 2.19, se puede observar los residuos obtenidos entre el modelo real (modelo simulado) y el modelo estimado para éste caso. Como una primera característica observable, está la poca variación respecto al plano constantemente

igual a cero por parte de los puntos. Esto permite intuir rápidamente que hubo un muy buen ajuste entre los datos estimados a partir del modelo presentado en la figura 2.18 y la superficie β_1 que se consideró para la simulación en 2.17. Más aún, se observa muy poca variabilidad entre los puntos residuales y al plano constantemente igual a cero. Esto ocurre, para los tres instantes de tiempo que fueron fijados, es decir, para $t = 1, t = 50$ y $t = 100$. Sin embargo, una mirada sobre los histogramas y los qq-plots de estos datos, nos permitirán saber si el comportamiento de estos residuales es bastante aproximado a la distribución normal.

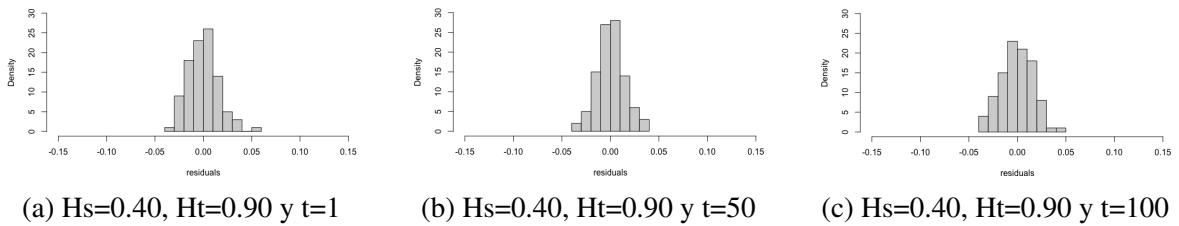


Figura 2.20 Histogramas de los Residuos: modelo $(\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}})$

En la figura 2.20 se presentan los histogramas para cada una de las distribuciones residuales en los instantes de tiempo $t = 1, t = 50$ y $t = 100$ del modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$. Cabe destacar que, estos histogramas presentan una similitud bastante buena con la distribución normal, es decir, no se muestra un comportamiento genérico de colas pesadas. Sin embargo, como ya lo hemos venido haciendo con el resto de los modelos, revisaremos a continuación los gráficos cuantil-cuantil entre la distribución normal teórica y la distribución de los residuales observada en cada histograma.

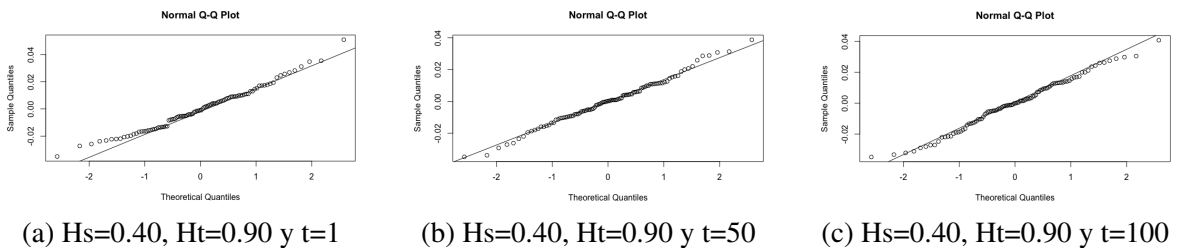


Figura 2.21 Gráficos Q-Q plot de los residuos para β_1

Para el conjunto de gráficos presentados en 2.21 podemos observar en cada gráfico cuantil-cuantil para los tiempos $t = 1, t = 50$ y $t = 100$ respectivamente, como la mayoría de los puntos tanto en el como en la centro de la diagonal están más cercanos a ésta, con excepción de uno cuantos que se escapan, sin embargo, en relación a los residuales anteriores y para el presente conjunto de datos residuales se aprecia un mejor ajuste.

Falta, mirar la gráfica de ECM para el estimador $\hat{\beta}_1$ de este caso, para evidenciar el comportamiento estable del estimador.

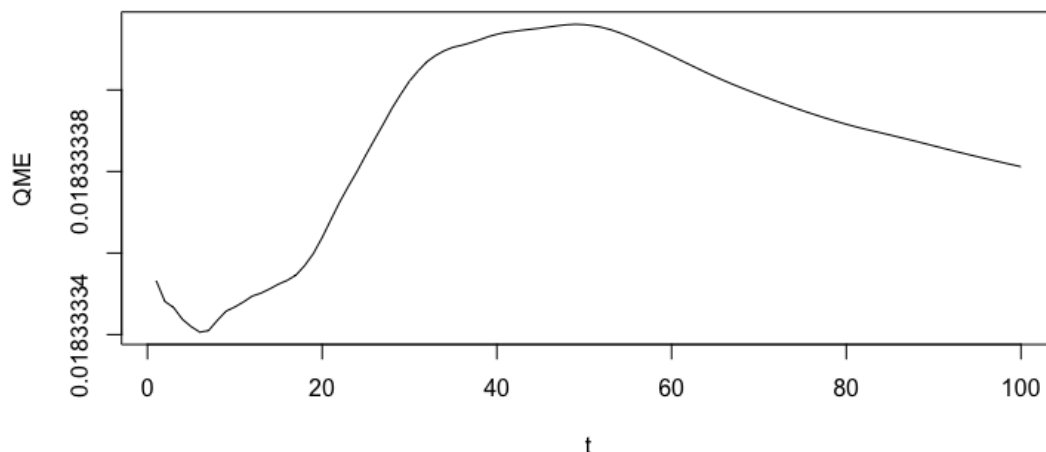


Figura 2.22 (ECM): modelo $\hat{Y} = \hat{\beta}_1 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$

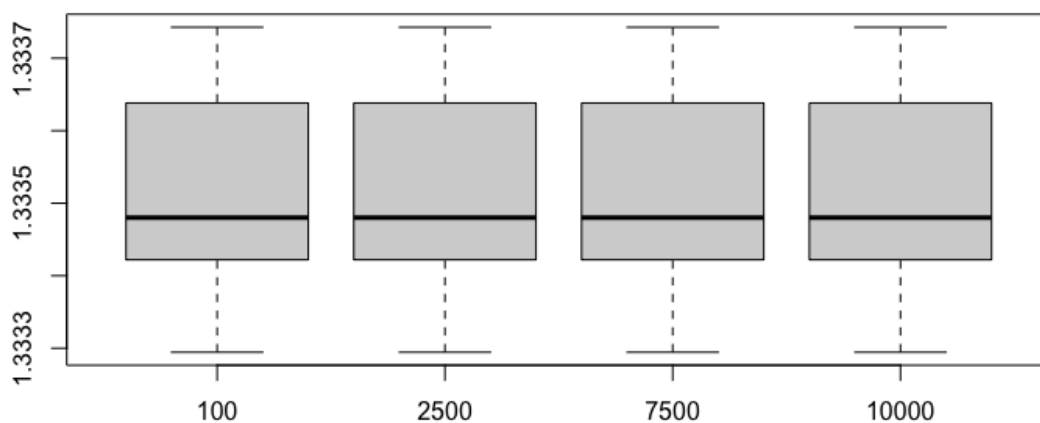


Figura 2.23 Box-plot para $\hat{\beta}_1$

A partir de la figura 2.22, correspondiente a la gráfica del error cuadrático medio ECM para $\hat{\beta}_1$ con índice de Hurst temporal $H_t = 0,90$, el cual se construyó bajo la misma lógica que se menciona para los modelos anteriores, podemos acotar que presente el mis comportamiento errático para los primero 60 tiempo t considerados y posteriormente se mantiene estable alrededor de valor muy pequeños que se muestran a la izquierda del gráfico. Por lo tanto, se evidencia un buen comportamiento asintótico para el estimador mencionado al inicio del párrafo. Más aún, afianzándonos en la figura 2.23 podemos asegurar el comportamiento estable de los datos durante las 10000 iteraciones realizadas ya que comparten rangos intercuantílicos exactamente iguales.

2.7.12. Simulación: modelo($Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$)

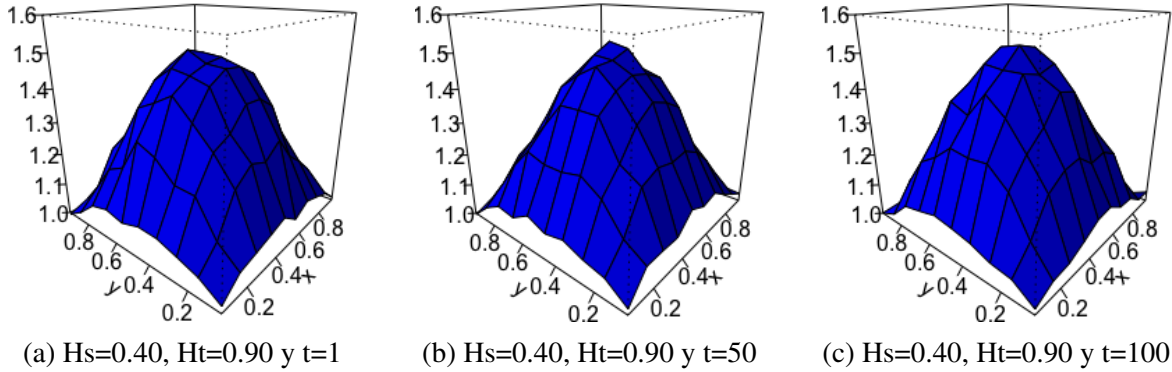


Figura 2.24 Modelo $Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$

A partir de la figura 2.24, se muestran las gráficas de β_2 para distintos t , las cuales son superficies curvas que se perturban mediante un ruido aditivo fraccionario coloreado, es decir, un ruido coloreado en espacio cuyo índice de Hurst considerado es $H_s = 0,40$ y fraccionario en tiempo con índice de Hurst $H_t = 0,90$. Cabe destacar, que las superficies para los tiempos escogidos $t = 1, t = 50$ y $t = 100$ son muy similares entre si, pues mantienen un comportamiento parecido en cuanto a la rugosidad del terreno simulado por estas.

Results of Global Regression				
Residuals:				
Mín	1Q	Median	3Q	Máx
-0,24	-0,14	-0,01	0,13	0,32
Coefficients:				
Estimate	Std. Error	t value	Pr(> t)	
1,33	0,001	983,3	< $2e^{-16}$	
Residual standard error: 0,154 on 9999 degrees of freedom				
Multiple R-squared: 0,984				
Adjusted R-squared: 0,9839				
F-statistic: $6,13e^5$ on 1 and 9999 DF, p-value: < $2,2e^{-16}$				
Extra Diagnostic information				
Residual sum of squares: 237,1161				
Sigma(hat): 0,1540011				
AIC: -9035,135				
AICc: -9035,134				
BIC: -19002,29				

Cuadro 2.7 Resultados de la regresión global $Y = \beta_2 * STAR(1, 1) + \epsilon_{H_{s4}, H_{t90}}$

De la regresión global presentada en el cuadro 2.7, se tiene que la significancia estadística para coeficiente β_2 que acompaña al modelo $Y = \beta_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$ es bastante buena. Es decir, bajo niveles de significancia para $\alpha = 0,001$, $\alpha = 0,01$, $\alpha = 0,05$, $\alpha = 0,1$ y $\alpha = 1$ el valor crítico $2e^{-16}$ es siempre menor a cualquiera de estos niveles mencionados. Por lo tanto, es pertinente considerarlo dentro de la regresión. Por otra parte, para la bondad de ajuste del modelo al igual que para la complejidad de su estructura, se consideran valores lo más bajo posibles de AIC y BCI , en este sentido, estos valores respectivamente son $AIC = -9035,135$ y $BCI = -19002,29$, es decir, el modelo simulado se ajusta bastante bien. Además, considerando el valor de R^2 múltiple que es igual a 0,984, podemos asegurar que el modelo predice aproximadamente el 98% de la variabilidad de los datos sobre la superficie. Sin embargo, para la suma de residuos al cuadrado al valor obtenido es 237,1161, el cual es bastante alto en relación al de modelos anteriores.

Results of Geographically Weighted Regression				
Model calibration information				
Kernel function: Gaussian				
Adaptive bandwidth: 147 (number of nearest neighbours.)				
Regression points: the same locations as observations are used.				
Distance metric: Euclidean distance metric is used.				
Summary of GWR coefficient estimates:				
Mín	1Q	Median	3Q	Máx
1,02	1,11	1,12	1,30	1,45
Diagnostic information				
Number of data points: 10000				
Effective number of parameters (2trace(S) - trace(S'S)): 28,03761				
Effective degrees of freedom (n-2trace(S) + trace(S'S)): 9971,962				
AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): -34000,38				
AIC (GWR book, Fotheringham, et al. 2002, GWR p. 96, eq. 4.22): -34021,71				
BIC (GWR book, Fotheringham, et al. 2002, GWR p. 61, eq. 2.34): -43863,77				
Residual sum of squares: 19,46016				
R-square value: 0,9184535				
Adjusted R-square value: 0,9182242				

Cuadro 2.8 Resultados de la regresión ponderada $Y = \beta_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$

El cuadro 2.8, nos presenta el resumen del análisis realizado sobre la regresión ponderada para el modelo $Y = \beta_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$. En este sentido, podemos evidenciar que los valores para AIC y BCI , los cuales respectivamente son $AIC = -34021,71$ y $BCI = -43863,77$ son significativamente más pequeños que los valores arrojados por la regresión global. Por otra parte, el valor del $R^2 = 0,9182242$ nos indica que el modelo es capaz de captar el aproximadamente el 92% de la variabilidad de los datos sobre la superficie estimada.

Más aún, la suma de residuos al cuadrado cuyo valor es 19,46016, presenta una diferencia con respecto al presentado en el análisis del modelo global significativamente más pequeño, es decir, para el modelo de regresión ponderado que considera el coeficiente de regresión dado por los β_2 , se concluye que las estimaciones echas por este son más precisas que las que pueda realizar el modelo global analizado previamente.

2.7.13. Estimación: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)

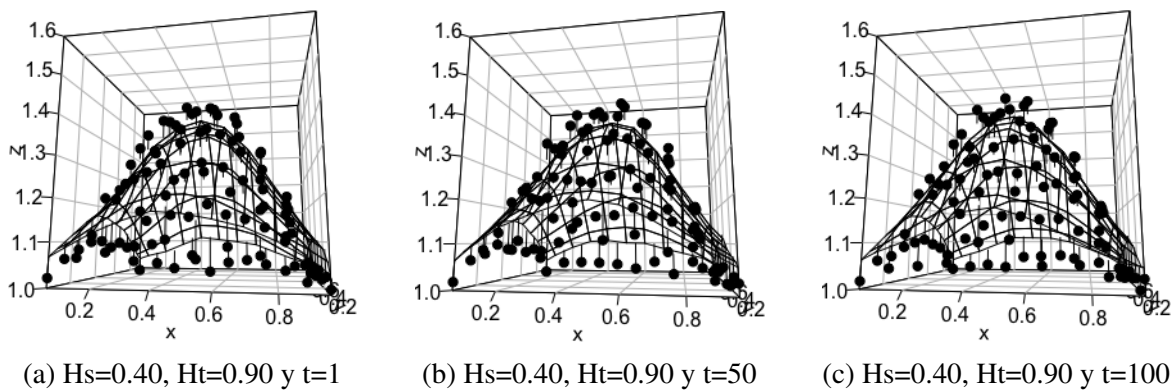


Figura 2.25 $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$

A partir de las superficies obtenidas durante las simulaciones anteriores, las cuales sirven como referencia de valores reales a ser estimados. Proponemos $\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$ como el estimador de las observaciones en la figura 2.25. Ahora bien, los puntos de color negro ubicados sobre la superficie curva levemente perturbada, son las observaciones estimadas sobre superficie β_2 en distintos instante de tiempo considerados. Ahora bien, para los tres casos dispuestos en la figura 2.25, es decir (a) con $t = 1$, (b) con $t = 50$ y (c) $t = 100$, las estimaciones son bastante similares entre si y muy cercanas a la superficie simulada. Esto, nos permite intuir que el modelo ponderado en relación a una bondad de ajuste es bastante bueno. Sin embargo, una mirada al gráfico de los residuos para el modelo ajustado, nos ayudara en la verificación que estos no presentan tanta variabilidad respecto a la superficie que estiman y se comportan aproximadamente normal.

2.7.14. Residuos: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)

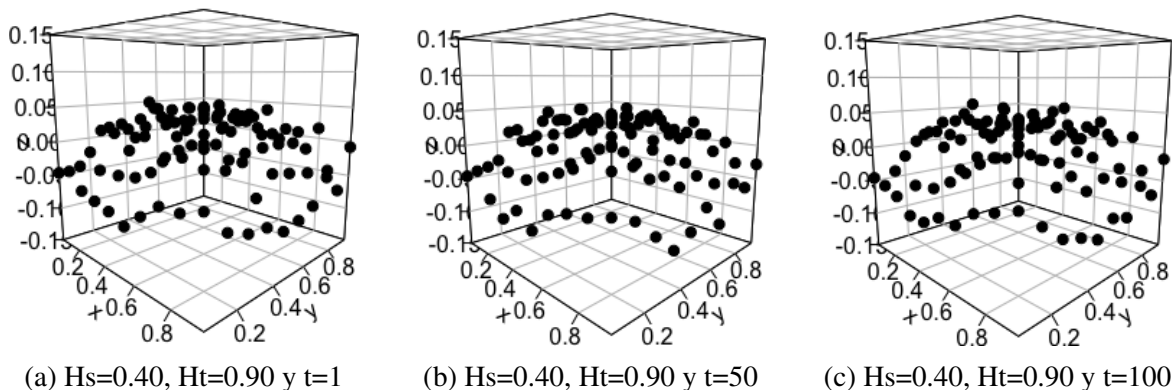


Figura 2.26 Residuos: modelo ($\hat{Y} = \hat{\beta}_2 * STAR(1, 1) + \varepsilon_{H_{s4}, H_{t90}}$)

La figura 2.26, muestra los residuos obtenidos entre el modelo real (modelo simulado) y el modelo estimado, los cuales presentan una variabilidad bastante alta para los tres instantes de tiempo que fueron fijados, es decir, para $t = 1, t = 50$ y $t = 100$. Cabe destacar, que éstas variaciones residuales ocurren alrededor del plano constantemente igual a cero, por lo que podríamos casi afirmar que los residuos tienen un comportamiento aproximadamente normal. Sin embargo, basándonos en análisis de los modelos anteriores podríamos afirmar que presentan problemas de ajuste a dicha distribución en las colas. Ahora bien, para ello realizaremos una revisión de los histogramas y los gráficos cuantil-cuantil relacionados a los datos residuales.

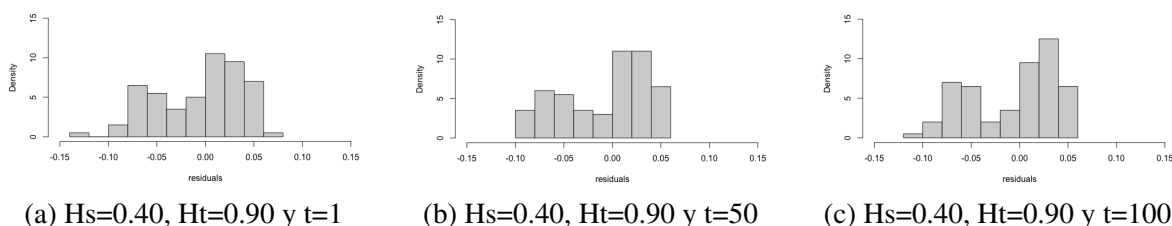


Figura 2.27 Histogramas de los residuos para β_1

Mediante la figura 2.27, es posible apreciar un comportamiento bimodal en los histogramas de las tres gráficas. Esto, podría arreglarse aumentando el número de observaciones y por ende el número de estimaciones en el modelo original. Sin embargo, este hecho no representa mayor problema dado que desde nuestro teorema principal 2.1 somos capaces de asumir un comportamiento estable para los datos en cierto sentido. Ahora bien, como un hecho visual, también se podría asumir que la forma de doble campana que presentan los

histogramas resulta de suma de distribuciones normales y que para grandes cantidades de datos estás tenderían aproximadamente a una distribución normal.

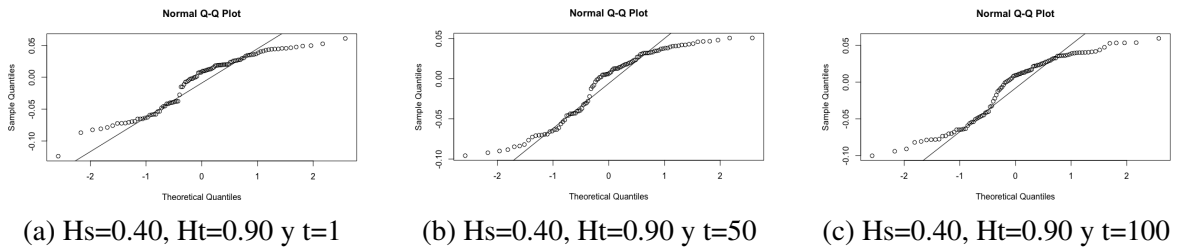


Figura 2.28 Gráficos Q-Q plot de los residuos para $\hat{\beta}_2$

La figura 2.28 nos muestra, los gráficos qq-plots para cada uno de los histogramas de datos residuales anteriores. Estos gráficos nos muestran que todas las distribuciones residuales de los t considerados presentan problemas de ajuste tanto en las colas como en el centro de la distribución. Sin embargo, los anterior puede deberse a que los datos residuales poseen un tipo de dependencia entre ellos que se puede expresar como suma de distribuciones normales. Por otro lado, es importante destacar que para los fines de nuestro trabajo, el uso del cálculo para el ECM juega un rol importante nuevamente, pues quien me asegura el buen comportamiento asintótico de mis distribuciones finito dimensionales. En ese sentido, continuaremos con la revisión de la gráfica que modela el comportamiento del ECM.

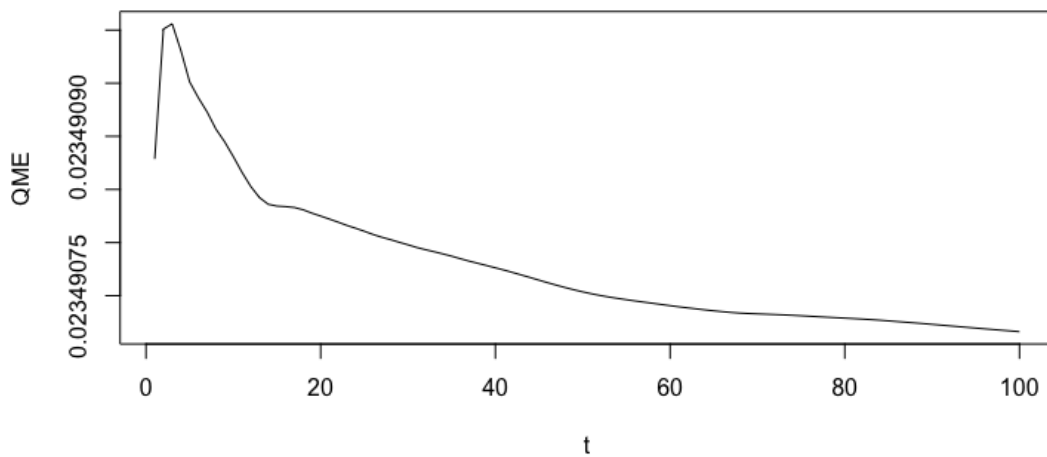
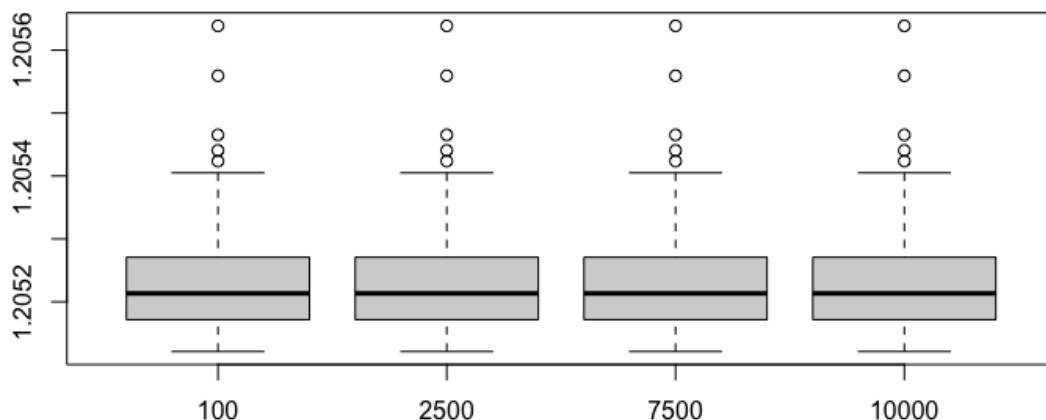


Figura 2.29 Error cuadrático medio para $\hat{\beta}_2$

Figura 2.30 Box-plot $\hat{\beta}_2$

Finalmente, desde la figura 2.29 podemos observar que el comportamiento del ECM en cuanto a sus iteraciones y para el estimador β_2 , presenta un comportamiento y errático para los primeros tiempos considerados. Sin embargo, posterior a ello sufre un decaimiento acelerado y tendiente a cero, lo que corrobora la estabilidad de dicho estimador. Por otra parte, el gráfico de caja y bigote nos muestra nuevamente el comportamiento igualitario a nivel intercuartilico, durante los periodos de datos temporales usados para las iteraciones. Así, podemos concluir que para los modelos propuestos dentro de las simulaciones se evidencia un mejor ajuste por parte de los modelos espacio-temporales ponderados sobre distintas superficies que presenta niveles de rugosidad presente que se rigen por procesos con larga memoria.

Capítulo 3

Trabajo futuros: (Estimación de parámetros en SPEDs asociadas a un modelo GTWAR)

Las ecuaciones en derivadas parciales estocásticas (SPDEs) con ruido fraccionario coloreado, han captado la atención de gran cantidad de investigadores en los últimos años. Esto, por el nivel de complejidad que muestran las estructuras de covarianza presentes en su componente espacial, sumado a la perturbación que poseen a nivel temporal. Sin embargo, aunque pocas son las referencias encontradas con la característica que se menciona sobre la estructura de covarianza espacio-temporal, es decir, con ruido fraccionario coloreado; en los trabajos Balan and Tudor (2007) y Torres et al. (2014) se explicitan modelos típicos a estudiar que dejan evidencia de la existencia de la solución y de la clase de ruido.

Ahora bien, nuestro principal interés es por el modelo para a la ecuación del calor estocástica en su forma lineal, con dominio sobre $(0, T) \times \mathbb{R}^d$, ya que estas poseen una estructura de covarianza espacial del ruido gaussiano que es fraccionaria en tiempo, con índice de Hurts $H \in (1/2, 1)$. Es decir, un ruido fraccionario coloreado. Por otra parte, en el artículo Gyöngy (1999), los autores estudian la discretización de las soluciones para la ecuación del calor dirigida por un ruido fraccionario en tiempo y coloreado en espacio.

Nosotros en el presente trabajo, haciendo uso de los resultados y técnicas mencionadas anteriormente, realizamos una discretización espacial para el operador de segundo orden en una base finita de autofunciones, con respecto a la variable espacial. Cabe destacar, que para ello utilizamos resultados de los trabajos en Torres et al. (2014) que permitieron la facil transición entre las expresiones que representan a dicho operador sobre base de autofunciones espaciales. Finalmente, aplicando un esquema de Euler sobre wel modelo semidiscretizado,

pero con respecto a la variable tiempo, obtuvimos un modelo discretizado completamente en espacio y tiempo, el cual al estar sobre una base finita de autofuncion representa un modelo autoregresivo espacial y temporalmente ponderado, es decir, un GTWAR.

La presente memoria, fue organizada en dos secciones. La primera sección 3.1, que contiene una revisión de los elementos y resultados más relevantes introducidos por los investigadores en Dalang (1999), Hörmander (2015), Balan and Tudor (2007), Torres et al. (2014), Delkhosh et al. (2012), Gyöngy (1999) y Boursès (2018) que aportaron a nuestro resultado asentando las bases sobre las cuales realizamos los cálculos para la discretización del modelo. En la segunda sección 3.2, realizamos la discretización de la ecuación del calor, mediante el procedimiento descrito, que nos permitió llegar al modelo GTWAR.

3.1. SPDEs con ruido fraccionario coloreado

A partir de los resultados planteado en Dalang (1999), las ecuaciones en derivadas parciales estocásticas (SPDEs) perturbadas por términos de ruido que tienen una estructura de covarianza espacial correlacionada o coloreada, y con ruido blanco en el tiempo, tiene la siguiente forma:

$$Lu = \dot{B} \quad (3.1)$$

donde L es un operador diferencial parcial, \dot{B} es un ruido gaussiano, que mayormente es blanco, pero posiblemente con algo de correlación a nivel espacial y u la solución débil de 3.1, que explícitamente se escriben como:

$$u(t, x) = \int_0^t \int_{\mathbb{R}^d} \Gamma(t-s, x-y) \dot{B}(ds, dy) \quad (3.2)$$

Sin embargo, la solución 3.2 es válida cuando las funciones de Green, (definidas en Delkhosh et al. (2012)), denotadas por Γ y asociadas con L , son una función y no una distribución.

Ahora bien, nuestro interés principal, está relacionado con la ecuación del calor estocástica, lineal y perturbada por un ruido coloreado fraccionario, la cual se define por la siguiente expresión:

$$\begin{aligned}
du(t,x) &= (\theta u_x x(t,x) + f(t,x)u(t,x))dt + \sigma dW^H(t,x) & (3.3) \\
u(t,0) &= u(t,1) = 0, \quad t > 0 \\
u(0,x) &= u_0(x), \quad x \in [0,1],
\end{aligned}$$

donde W^H es un movimiento Browniano fraccionario coloreado en espacio y en tiempo, definido en un espacio de probabilidad (Ω, F, \mathbb{P}) , u_0 es una función continua en $[0, 1]$ tal que $u_0(0) = u_0(1) = 0$, θ y σ constantes positivas, y f una función que depende de las variables espacio-temporal y satisface ciertas condiciones de regularidad de tipo Lipchitz para asegurar la existencia de soluciones de la ecuación, es decir f satisface

$$|f(t,x) - f(t,y)| \leq |x - y| \quad \forall t \in [0, T], x, y \in [0, 1]. \quad (3.4)$$

Por otra parte, la solución fundamental asociada a la ecuación del calor planteada en 3.3, está planteada en los trabajos de Dalang (1999), más aún se define la extensión de la medida de martingala integral estocástica con aplicaciones a SPDEs espacialmente homogéneas. Este hecho, permitió plantear soluciones de ecuaciones del tipo 3.1, donde la función de Green denotada por Γ y que se define en Delkhosh et al. (2012) como la solución fundamental de la ecuación del calor, no es una función propiamente sino una distribución de Schwartz.

Particularmente, para la ecuación 3.3 en dimensiones $d = \{1, 2\}$ se utiliza una Γ , dada por el kernel de Riesz para representar el ruido espacial, ya que temporalmente, como se mencionó al inicio, mantiene un comportamiento de ruido blanco, que representa un proceso Gaussiano de media cero y función de covarianza

$$E(B_t^H, B_s^H) = \frac{1}{2}(s^{2H} + t^{2H} - |t - s|^{2H}) \quad 0 < H < 1, \quad (3.5)$$

lo cual, es un movimiento Browniano fraccionario.

Observación 3.1. *Para tener un mejor conocimiento sobre la definición de un espacio de Schwartz o espacio de las funciones de decrecimiento rápido, se recomienda revisar el apartado 3.4.6. de la siguiente literatura Bourlès (2018). Además, recordemos que las funciones de Green, también llamadas soluciones fundamentales en Delkhosh et al. (2012), son funciones kernel, tal como se definen en Hörmander (2015). Estas funciones, son usadas dentro del operador lineal integral, permitiendo encontrar las soluciones de ecuaciones diferenciales. En este sentido, cuando se habla de un ruido coloreado fraccionario, se hace mención a la estructura de las perturbaciones espaciales que mantienen un comportamiento gaussiano en el tiempo, es decir, un comportamiento típicamente de ruido blanco en el tiempo,*

pero posiblemente con alguna estructura de correlación espacial, tal como se menciona en los trabajos de Dalang (1999). Particularmente, para el caso de la ecuación del calor, la función kernel de Riesz se usa habitualmente para representar el ruido espacial, en dimensiones $d = \{1, 2\}$. Este ejemplo, se puede encontrar hecho en Dalang (1999).

Por otra parte, desde los aportes hechos por los investigadores en el trabajo Gyöngy (1999), se introduce un importante resultado que permite realizar la discretización de las soluciones en la ecuación 3.3 usando el sistema de ecuaciones diferenciales que se muestra a continuación:

$$du^M(t, x_m) = \theta M^2 (u^M(t, x_{m+1}) - 2u^M(t, x_m) + u^M(t, x_{m-1}))dt + f(t, x_m)u^M(t, x_m)dt + \sigma d(W^H(t, x_{m+1}) - W^H(t, x_m)). \quad (3.6)$$

para,

$$u^M(t, 0) = u^M(t, 1) = 0 \quad (3.7)$$

$$u^M(0, x_m) = u_0(x_m), \quad m = 1, \dots, m-1 \quad (3.8)$$

donde $M \geq 1$, $x_m = \frac{m}{M}$ son los puntos de la malla para $m = 0, \dots, M$, con incrementos definidos por $\Delta x = \frac{1}{M}$, $u^M(t, \cdot)$ un proceso solución para sistema de ecuaciones con $m = 1, \dots, m-1$, u_0 es una función continua en $[0, 1]$ tal que $u_0(0) = u_0(1) = 0$, θ y σ constantes positivas, f una función, tal que depende de las variables espacio-temporal, que satisface ciertas condiciones de regularidad de tipo Lipchitz y W^H es un movimiento Browniano fraccionario coloreado en espacio.

3.2. Discretización de soluciones

Sea $M \geq 1$ un entero. Definimos los puntos de la malla $x_m = \frac{m}{M}$ para $m = 0, \dots, M$; y el tamaño de la malla $\Delta x = \frac{1}{M}$. Ahora, usamos el esquema estándar de diferencias finitas para la aproximación espacial de 3.3 desde Gyöngy (1999). Sea el proceso $u^M(t, \cdot)$. Definimos como solución del sistema de ecuaciones diferenciales estocásticas (para $m = 1, \dots, M-1$)

$$du^M(t, x_m) = \theta M^2 (u^M(t, x_{m+1}) - 2u^M(t, x_m) + u^M(t, x_{m-1}))dt + f(t, x_m)u^M(t, x_m)dt + \sigma d(W^H(t, x_{m+1}) - W^H(t, x_m)). \quad (3.9)$$

para,

$$u^M(t, 0) = u^M(t, 1) = 0$$

$$u^M(0, x_m) = u_0(x_m), \quad m = 1, \dots, m-1$$

Denotamos $u_m^M(t) := u^M(t, x_m)$ y $W_m^{H,M}(t) := \sqrt{M}[W^H(t, x_{m+1}) - W^H(t, x_m)]$ para $m = 1, \dots, M$ y escribimos el sistema 3.3 como

$$du_m^M(t) = \theta M^2 \sum_{i=1}^{M+1} D_{mi} u_i^M dt + f(t, x_m) u_m^M(t) dt + \sqrt{M} \sigma dW_m^{H,M}(t)$$

con valor inicial $u_m^M(0) = u_0(x_m)$, para $m = 1, \dots, M-1$, donde $D = (D_{mi})_{m,i}$ es una matriz cuadrada de tamaño $M-1$, con elementos $D_{mm} = -2, D_{mi} = 1$ para $|m-i| = 1$, $D_{mi} = 0$ para $|m-i| > 1$.

Sea $A = M^2 D$. Usando el método de variación de constantes, la solución exacta al sistema 3.3 se obtiene

$$\begin{aligned} u^M(t, x_m) &= \frac{1}{M} \sum_{l=1}^{M-1} \sum_{j=1}^{M-1} \exp(\lambda_j^M \theta t) \varphi_j(x_m) \varphi_j(x_l) u_0(x_l) \\ &+ \int_0^t \frac{1}{M} \sum_{l=1}^{M-1} \sum_{j=1}^{M-1} \exp(\lambda_j^M \theta (t-s)) \varphi_j(x_m) \varphi_j(x_l) f(s, x_l) u^M(s, x_l) ds \\ &+ \int_0^t \frac{1}{M} \sum_{l=1}^{M-1} \sum_{j=1}^{M-1} \exp(\lambda_j^M \theta (t-s)) \varphi_j(x_m) \varphi_j(x_l) \sigma dW_l^{H,M}(s) \end{aligned} \quad (3.10)$$

donde λ^M, φ_j son los autovalores y autofunciones de A .

Ahora bien, para introducir la discretización en tiempo para el problema semidiscretizado en espacio consideramos un entero $N \geq 1$ y algún tiempo final $T > 0$. Sea $\Delta t = \frac{T}{N}$ y definimos los tiempos discretos $t_n = n\Delta t$ para $n = 0, \dots, N$. Ahora consideremos la ecuación 3.3 en el intervalo de tiempo $[t_n, t_{n+1}]$ como sigue:

$$u^M(t_{n+1}) = e^{\theta A \Delta t} u^M(t_n) + \int_{t_n}^{t_{n+1}} e^{\theta A (t_{n+1}-s)} f(s, x^M) u^M(s) ds + \int_{t_n}^{t_{n+1}} e^{\theta A (t_{n+1}-s)} \sqrt{M} \sigma dW^{H,M}_s \quad (3.11)$$

con $u_m^M(t) = (u_m^M(t), m = 1, \dots, M-1)$, $u_m^M(t) = u^M(t, x_m)$, el vector $f(s, x^M) = (f(s, x_m), m = 1, \dots, M-1)$ y $dW^{H,M}(s) = (dW^{H,M}(s, x_m), m = 1, \dots, M+1)$.

Usando un esquema de discretización de Euler en la malla de tiempo dada y denotando

$$u^{M,N}(t_{n+1}) = e^{\theta A \Delta t} u^{M,N} t_n + e^{\theta A \Delta t} f(t_n, x^M) u^{M,N} t_n \Delta t + e^{\theta A \Delta t} \sqrt{m} \sigma [W^{H,M}(t_{n+1}) - W^{H,M}(t_n)],$$

obtenemos la aproximación $U_m^{n+1} \approx u(t_{n+1}, x_m)$ dado por

$$\begin{aligned}
 U_m^{n+1} &= \frac{1}{M} \sum_{l=1}^{M-1} \sum_{j=1}^{M-1} \exp(\lambda_j^M \theta \Delta t) \phi_j(x_m) \phi_j(x_l) U_l^n \\
 &+ \Delta t \frac{1}{M} \sum_{l=1}^{M-1} \sum_{j=1}^{M-1} \exp(\lambda_j^M \theta \Delta t) \phi_j(x_m) \phi_j(x_l) f(t_n, x_l) U_l^n \quad (3.12) \\
 &+ \frac{1}{M} \sum_{l=1}^{M-1} \sum_{j=1}^{M-1} \exp(\lambda_j^M \theta \Delta t) \phi_j(x_m) \phi_j(x_l) \sigma [W_l^{H,M}(t_{n+1}) - W_l^{H,M}(t_n)]
 \end{aligned}$$

donde la ecuación 3.12 puede expresarse como un modelo GTWAR. Sin embargo, para ello se debe considerar dos puntos importantes los cuales desprenden dos problemas importantes que permiten la continuidad a la presente investigación:

- La estimación del modelo GTWAR considerando un ruido coloreado fraccionario en la estructura de covarianza del ruido.
- La identificabilidad de los coeficientes entre el modelo GTWAR con ruido coloreado fraccionario con los coeficientes del modelo 3.12

Apéndice A

El detalle de nuestro principal resultado, el cual se expone en el capítulo 2 se deja explícito en la siguiente dirección:

<https://drive.google.com/file/d/1JhyFfN9m-KfUI0wBLkIEHrvC2guw4ar0/view>

Bibliografía

- E. Azmoodeh, T. Sottinen, L. Viitasaari, and A. Yazigi. Necessary and sufficient conditions for hölder continuity of gaussian processes. *Statistics and Probability Letters*, 94:230–235, 2014. ISSN 0167-7152. doi: <https://doi.org/10.1016/j.spl.2014.07.030>. URL <https://www.sciencedirect.com/science/article/pii/S0167715214002776>.
- R. Balan and C. Tudor. The stochastic heat equation with a fractional-colored noise: Existence of the solution. 2007.
- F. Biagini, Y. Hu, B. Øksendal, and T. Zhang. *Stochastic Calculus for Fractional Brownian Motion and Applications*. Probability and Its Applications. Springer London, 2008. ISBN 9781846287978. URL <https://books.google.com.gi/books?id=ccdWU2NW9f4C>.
- H. Bouchlé. 3 - topological vector spaces. In H. Bouchlé, editor, *Fundamentals of Advanced Mathematics 2*, pages 117–196. Elsevier, 2018. ISBN 978-1-78548-249-6. doi: <https://doi.org/10.1016/B978-1-78548-249-6.50003-1>. URL <https://www.sciencedirect.com/science/article/pii/B9781785482496500031>.
- L. L. Cam. The Central Limit Theorem Around 1935. *Statistical Science*, 1(1):78 – 91, 1986. doi: 10.1214/ss/1177013818. URL <https://doi.org/10.1214/ss/1177013818>.
- A. Can and I. Megbolugbe. Spatial dependence and house price index construction. *The Journal of Real Estate Finance and Economics*, 14(1-2):203–22, 1997.
- J. Chen, S.-L. Shaw, H. Yu, F. Lu, Y. Chai, and Q. Jia. Exploratory data analysis of activity diary data: a space–time gis approach. *Journal of Transport Geography*, 19(3):394–404, 2011.
- T. Cheng and D. Williams. Space-time analysis of crime patterns in central london. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 39:47–52, 2012.
- N. Cressie. *Statistics for spatial data*. 1993.
- R. C. Dalang. Extending martingale measure stochastic integral with applications to spatially homogeneous s. p. d. e’s. *Electronic Journal of Probability [electronic only]*, 4:Paper No. 6, 29 p.–Paper No. 6, 29 p., 1999. URL <http://eudml.org/doc/119958>.
- M. Delkhosh, M. Delkhosh, and M. Jamali. Green’s function and its applications. *Journal of Basic and Applied Scientific Research*, 2012:8865–8876, 09 2012.

- H. Djalihu, A. Djuraidah, and A. M. Soleh. Mixed geographically and temporally weighted autoregressive to modeling the levels of poverty population in java in 2012-2018. *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, 54:55–67, 11 2020.
- K. Ellegård and U. Svedin. Torsten hägerstrand’s time-geography as the cradle of the activity approach in transport geography. *Journal of Transport Geography*, 23:17–25, 07 2012. doi: 10.1016/j.jtrangeo.2012.03.023.
- Fermín, Torres, Gomez, Roa, and Araya. Estimation of a weighted regression model with fractional-colored noise: Consistency. *arXiv(Preprint)*, 2023.
- A. Fotheringham, C. Brunsdon, and M. Charlton. Geographically weighted regression: The analysis of spatially varying relationships. *John Wiley & Sons*, 13, 01 2002.
- A. Fotheringham, R. Crespo, and J. Yao. Geographical and temporal weighted regression (gtwr). *Geographical Analysis*, 47, 04 2015a.
- A. S. Fotheringham, R. Crespo, and J. Yao. Geographical and temporal weighted regression (gtwr). *Geographical Analysis*, 47(4):431–452, 2015b.
- I. Gyöngy. Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise ii. *Potential Analysis*, 11(1):1–37, 1999.
- Harianto, N. Waego, and E. Sumarminingsih. Geographically and temporally weighted regression model with gaussian kernel weighted function and bisquare kernel weighted function. *IOP Conference Series: Materials Science and Engineering*, 1115:012063, 03 2021.
- L. Hörmander. *The Analysis of Linear Partial Differential Operators I: Distribution Theory and Fourier Analysis*. Classics in Mathematics. Springer Berlin Heidelberg, 2015. ISBN 9783642614972. URL <https://books.google.fr/books?id=aaLrCAAAQBAJ>.
- C. M. Hurvich, J. S. Simonoff, and C.-L. Tsai. Smoothing parameter selection in nonparametric regression using an improved akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60, 1998.
- A. Kolmogorov. Wienersche spiralen und einige andere interessante kurven im hilbertschen raum. *Dokl. Acad. Sci. URSS*, 26, 01 1940.
- X. Luo, Y. Chen, Z. Wang, H. Li, and P. Yidong. Spatial downscaling of modis land surface temperature based on a geographically and temporally weighted autoregressive model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, PP: 1–1, 07 2021. doi: 10.1109/JSTARS.2021.3094184.
- B. B. Mandelbrot and J. W. van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10:422–437, 1968.
- A. D. Miftahus Sholihin, Agus Mohamad Soleh. Geographically and temporally weighted regression (gtwr) for modeling economic growth using r. *International Journal of Computer Science and Network - IJCSN*, 10:800–805, 12 2017.

- I. S. Mishura, I. S. Mishura, J. S. Mišura, Y. Mishura, and Ū. S. Mišura. *Stochastic calculus for fractional Brownian motion and related processes*, volume 1929. Springer Science & Business Media, 2008.
- I. Nourdin. *Selected Aspects of Fractional Brownian Motion*, volume 4 of *Bocconi & Springer Series*. 2012. URL <https://hal.archives-ouvertes.fr/hal-01314412>.
- A. Nualart David Rascanu. Differential equations driven by fractional brownian motion. *Collectanea Mathematica*, 53(1):55–81, 2002. URL <http://eudml.org/doc/42846>.
- Y. Peng, W. Li, X. Luo, and H. Li. A geographically and temporally weighted regression model for spatial downscaling of modis land surface temperatures over urban heterogeneous regions. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7):5012–5027, 2019.
- X. Que, X. Ma, C. Ma, and Q. Chen. A spatiotemporal weighted regression model (stwr v1.0) for analyzing local nonstationarity in space and time. *Geoscientific Model Development*, 13(12):6149–6164, 2020. doi: 10.5194/gmd-13-6149-2020. URL <https://gmd.copernicus.org/articles/13/6149/2020/>.
- J. Shim and C. Hwang. Kernel-based geographically and temporally weighted autoregressive model for house price estimation. *PLOS ONE*, 13:e0205063, 10 2018. doi: 10.1371/journal.pone.0205063.
- K. Takahashi, M. Kulldorff, T. Tango, and K. Yih. A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. *International Journal of Health Geographics*, 7(1):1–14, 2008.
- S. Torres, C. A. Tudor, and F. Viens. Quadratic variations for the fractional-colored stochastic heat equation. *Electronic Journal of Probability*, 19:1–51, 2014.
- B. Wu, R. Li, and B. Huang. A geographically and temporally weighted autoregressive model with application to housing prices. *International Journal of Geographical Information Science*, 28:1186–1204, 05 2014.

