

CLUSTERIZACIÓN DE BARRIOS EN EL AGLOMERADO JUJUY - PALPALÁ POR VARIABILIDAD DE PRECIOS MINORISTAS DE FRUTAS Y VERDURAS

NEIGHBORHOOD CLUSTERING IN THE JUJUY - PALPALÁ AGGLOMERATION BASED ON VARIABILITY IN RETAIL PRICES OF FRUITS AND VEGETABLES

Solís, J. M.^{1*}, Carrasco, S. C.¹, Leaño, M. C.¹, Martínez, F. N.², Chaile, V. A.², Quespía Mamani, M. A.², Tolay Toconás, E. A.², Rosales, Y. B.², Chávez, J. L.² y Yarbe, G. I.²

¹ Cátedra de Bioestadística y Diseño Experimental

² Estudiantes avanzados de la carrera Ingeniería Agronómica de la FCA UNJu

*Autor para correspondencia:
juanmasolis@fca.unju.edu.ar

Período de Publicación:
Diciembre 2025

Historial:
Recibido: 02-10-25
Aceptado: 25-11-25

RESUMEN

En este trabajo se analiza la estructura espacial de la variabilidad de los precios minoristas de hortalizas y frutas para consumo en el aglomerado San Salvador de Jujuy-Palpalá. Se parte del supuesto de que, en mercados atomizados e informales como el hortícola, la dispersión de precios es reflejo de asimetrías de información y condiciones locales heterogéneas. A partir de datos relevados en distintos puntos de venta, se realizó un análisis de Procrustes Generalizado (GPA), se construyeron Modelos de Efectos Aleatorios y de conglomerados jerárquicos, y se aplicaron técnicas de autocorrelación espacial (Moran Global y Local). Los resultados indican que el factor *barrio* explica en promedio un 15 % de la variabilidad total, mientras que el *tipo de comercio* explica menos del 2 %. El análisis multivariado reveló la existencia de dos agrupamientos territoriales diferenciados, uno concentrado en los barrios del norte y otro en la zona sur del aglomerado. La significancia estadística de estos patrones fue confirmada mediante el índice de Moran Global ($I = 0,53$, $p = 0,003$), y complementada con la identificación de clústeres locales mediante la prueba LISA. Los hallazgos respaldan la hipótesis de una estructura espacial no aleatoria de precios, lo que sugiere la necesidad de incorporar herramientas analíticas espaciales en los estudios de mercados minoristas. Este enfoque ofrece una base para el diseño de políticas que busquen mejorar el acceso equitativo a alimentos en contextos urbanos complejos.

Palabras clave: análisis de conglomerados jerárquicos, autocorrelación espacial, variabilidad espacial

SUMMARY

This study analyzes the spatial structure of price variability for retail fruits and vegetables in the San Salvador de Jujuy-Palpalá urban agglomeration.

It assumes that, in an atomized and informal markets such as the horticultural sector, price dispersion reflects information asymmetries and heterogeneous local conditions. Using data collected from various retail outlets, a Generalized Procrustes Analysis (GPA) was conducted, along with the construction of Random Effects Models and hierarchical clustering, and the application of spatial autocorrelation techniques (Global and Local Moran's I). The results show that the *neighbourhood* factor accounts for an average of 15% of total variability, whereas the type of *retail outlet* explains less than 2%. Multivariate analysis revealed the existence of two distinct territorial clusters—one concentrated in the northern neighbourhoods and the other in the southern part of the metropolitan area. The statistical significance of these spatial patterns was confirmed through the Global Moran's I index ($I = 0.53$, $p = 0.003$), and further supported by the identification of local clusters using the LISA test. The findings support the hypothesis of a non-random spatial structure of prices, highlighting the relevance of incorporating spatial analytical tools in retail markets study. This approach provides a foundation for policy design aimed at improving equitable access to food in complex urban settings.

Keywords: hierarchical cluster analysis, spatial autocorrelation, spatial variability

INTRODUCCIÓN

Los precios constituyen uno de los principales indicadores de funcionamiento de un mercado o una economía (Market Information Organization of the Americas, Washington, D. C. (United States of America) & Instituto Interamericano de Cooperación para la Agricultura (IICA), 2016). Las oscilaciones de los precios de los productos básicos agrícolas son indicadores de cambios en su oferta y demanda. Contar con información transparente y oportuna sobre los mercados es fundamental para la adopción de decisiones basadas en datos objetivos y para adoptar estrategias de seguridad alimentaria, la cual implica cuatro dimensiones: a) disponibilidad física de los alimentos, b) acceso económico y físico, c) empleo de los alimentos y c) estabilidad en el tiempo de las tres anteriores (FAO, 2023).

El sector hortícola de la provincia de Jujuy se caracteriza por una atomización de productores enmarcados en gran parte en la figura de agricultura familiar. Sumado a ello, la tenencia de la tierra corresponde en gran parte a agricultores arrendatarios, medieros o en sociedad con el propietario de la tierra. Según el Censo Nacional Agropecuario 2018 (CNA18), la superficie de parcelas productivas correspondiente a tierras en arrendamiento representa un 19% y un 3% a nivel país y provincial respectivamente, mientras que este porcentaje asciende a un 28% en el departamento de El Carmen, donde se concentra la mayor cantidad de producción y mercados de frescos de la provincia (INDEC, 2021). Estas figuras socio – productivas favorecen la existencia de un alto grado de informalidad en los productores y de sus trabajadores (Martínez et al., 2010; MAGYP, 2022), con escasa o nula información acerca de los costos de producción y de toda la cadena de comercialización. Por lo tanto, la formación de precios mayoristas y minoristas toma características oligopólicas, donde la presencia de intermediarios en la cadena de valor de las hortalizas juega un rol destacado.

Una de las causas de la dispersión de precios es la existencia de información imperfecta (Stigler, 1961; Stahl, 1996; Sorensen, 2000; Vicentin Masaro & Chara, 2020), y puede distinguirse entre una dispersión

de precios espacial y una temporal (Varian, 1981). La primera ocurre cuando las firmas¹ fijan precios en distintos momentos del tiempo, pero la posición de la firma en la distribución de precios general no cambia (se mantiene en la misma posición en el ranking de precios bajos/altos); en cambio, la segunda se observa cuando una misma firma ofrece precios de manera impredecible en el tiempo, variando el lugar en el ranking (Vicentin Masaro & Chara, 2020). Diversas publicaciones han explorado diferentes enfoques para analizar dispersión de precios minoristas en alimentos con datos georreferenciados (Perez, 2010; Vicentin Masaro & Chara, 2020; Masaro *et al.*, 2020) , pero no se han encontrado trabajos que describan la variabilidad de precios minoristas en mercados de frutas y verduras para consumo en fresco, al menos en la región NOA y particularmente la provincia de Jujuy.

Ante la escasez de información sobre el comportamiento de los precios en el eslabón minorista de la cadena de comercialización de frutas y hortalizas en el aglomerado Jujuy - Palpalá, y considerando que este análisis constituye un punto de partida ineludible para la caracterización de otros eslabones, este trabajo plantea los siguientes objetivos:

- Estimar la contribución relativa de los factores Barrio y Tipo de comercio sobre la variabilidad observada en la formación de precios minoristas de una canasta seleccionada de frutas y verduras, en el Aglomerado Gran Jujuy.
- Identificar y delimitar conglomerados o *clústeres* espaciales de barrios que presenten una distribución homogénea en sus precios relativos.

MATERIALES Y MÉTODOS

Se realizó un relevamiento mensual de los precios minoristas de los siguientes productos hortícolas y frutícolas destinados al consumo en fresco (en adelante, “productos”): papa *Solanum tuberosum* (kg), banana no importada grande *Musa paradisiaca* (docena), cebolla grande *Allium cepa* (kg), lechuga tipo crespa *Lactuca sativa* var. *crispa* (kg), naranja dulce para jugo *Citrus sinensis* (docena), manzana deliciosa *Malus domestica* (kg) y tomate perita grande *Solanum lycopersicum* (kg).

Se realizaron siete relevamientos de precios en los meses de julio, agosto, octubre y noviembre de 2023, y en marzo, abril y mayo de 2024, utilizando el método de encuesta directa. Los barrios relevados fueron los siguientes (Tabla 1):

Tabla 1: Barrios relevados durante el proyecto

Id	Barrio	Distrito	Id	Barrio	Distrito
1	Alto Comedero (A.C.)	Alto Comedero	10	Malvinas Argentinas (M.A.)	Este
2	Dr. Marcelino Vargas (D.M.V.)	Norte	11	Los Huaicos (L.H.)	Oeste
3	Ciudad de Nieva (C.D.N.)	Oeste	12	Bajo la Viña (B.L.V.)	Norte
4	Coronel Arias (C.A.)	Sur	13	Almirante Brown (A.B.)	Este
5	Los Perales (L.P.)	Norte	14	Alto Gorriti (A.G.)	Sur
6	Chijra (Ch.)	Norte	15	Cuyaya (Cu.)	Sur

¹ Nota: en este trabajo se referirá a *puntos de venta* en lugar de firmas.

Id	Barrio	Distrito	Id	Barrio	Distrito
7	San Pedrito (S.P.)	Este	16	San salvador de Velazco en el Valle de Jujuy (Centro)	Centro
8	Alto la Viña (A.L.V.)	Norte	17	Bajo Gorriti (B.G.)	Centro
9	Mariano Moreno (M.M.)	Sur	18	Palpalá (P.) ^{<?>}	Palpalá

Cada relevamiento se realizó durante la tercera semana de cada mes, con una extensión no mayor a los 10 días entre la primera y la última encuesta. La distribución de los puntos muestrales se estableció mediante un esquema de muestreo estratificado por conglomerados en una etapa para garantizar la representatividad de los distintos barrios (estratos), donde por razones de accesibilidad y seguridad, ya que los relevamientos fueron realizados por estudiantes de la carrera de Ingeniería Agronómica de la Facultad de Ciencias Agrarias, se decidió trabajar con un marco muestral reducido donde las unidades primarias de muestreo (UPM) estuvieron constituidas por todas las manzanas tangentes por alguna de sus cuadras al sistema de *avenidas del aglomerado*, y donde las unidades secundarias de muestreo (USM) fueron todos los puntos de venta localizados en cada UPM. Esta decisión metodológica, si bien implica una limitación en la representatividad espacial completa del aglomerado, no compromete la validez interna de los análisis realizados. Ello se debe a que el objetivo del estudio no consiste en inferir parámetros poblacionales absolutos -lo cual requeriría del empleo de factores de expansión desde el punto de vista de una estimación basada en el diseño-, sino en estimar la estructura relativa de la variabilidad de precios y explorar patrones de correlación espacial dentro del marco de observación definido, a través de una estimación basada en modelos (*model-based inference*), en la que las variables que definen los estratos están incluidas como variables explicativas en el modelo. En este sentido, el diseño adoptado resulta útil para cumplir con los fines exploratorios y analíticos del trabajo, con una base empírica consistente para identificar regularidades espaciales de interés.

En total se visitaron 270 puntos de muestreo, distribuidos proporcionalmente al número de viviendas según el Pre Censo de Viviendas Particulares y Colectivas (INDEC, 2020). De ese total, en cinco puntos de venta los vendedores no aceptaron participar y en otros 17 no se encontraron puestos de venta, lo que obligó a la relocalización de esos puntos de muestra. Al finalizar el relevamiento se obtuvieron 248 respuestas válidas. Cada punto de venta se clasificó según el tipo de comercio como almacén, mercado, mini mercado o *mini market*, supermercado, venta ambulante y verdulería (que también incluyó puestos en ferias minoristas). En las Tablas 2 y 3 puede observarse la distribución de respuestas *positivas* por barrio y por tipo de comercio respectivamente:

Tabla 2: Distribución de USM relevadas con respuestas positivas, por barrio según momento de muestreo

Barrio	Muestra						
	1	2	3	4	5	6	7
Almirante Brown	3	3	1	2	4	6	5
Alto Comedero	10	8	11	9	11	10	5
Alto Gorriti	2	3	0	1	4	4	4
Bajo Gorriti	1	0	0	0	2	2	1
Bajo La Viña	1	1	1	1	1	1	1
Chijra	2	2	3	2	2	2	2

Barrio	Muestra						
	1	2	3	4	5	6	7
Ciudad De Nieva	2	1	1	1	1	1	1
Coronel Arias	2	1	1	1	1	1	0
Cuyaya	3	4	0	2	1	3	3
Dr. Marcelino Vargas	1	1	1	1	0	0	0
Los Huaicos	2	1	1	1	1	1	1
Los Perales	1	1	1	2	1	0	0
Malvinas Argentinas	2	5	2	2	1	1	0
Mariano Moreno	2	1	1	2	1	1	0
San Pedrito	6	3	3	2	1	1	1
San Salvador De Velazco En El Valle De Jujuy	1	0	0	0	3	4	4
Alto La Viña	0	0	1	1	0	0	0
Palpalá	0	0	0	0	5	4	4
Total	41	35	28	30	40	42	32

Tabla 3: Distribución de USM relevadas con respuestas positivas, por tipo de comercio según momento de muestreo

Tipo de comercio	Muestra						
	1	2	3	4	5	6	7
Almacén	20	15	15	18	6	12	6
Mercado	5	6	1	2	11	15	10
Mini mercado	6	1	4	3	9	5	6
Supermercado	1	0	0	0	2	1	2
Venta ambulante	2	3	1	2	3	4	2
Verdulería	6	10	7	5	9	5	6
Otro	1	0	0	0	0	0	0
Total	41	35	28	30	40	42	32

En la Figura 1 se muestra la distribución espacial de las USM con respuesta positiva a lo largo de todo el relevamiento de datos, agrupadas con el fin de facilitar su lectura:

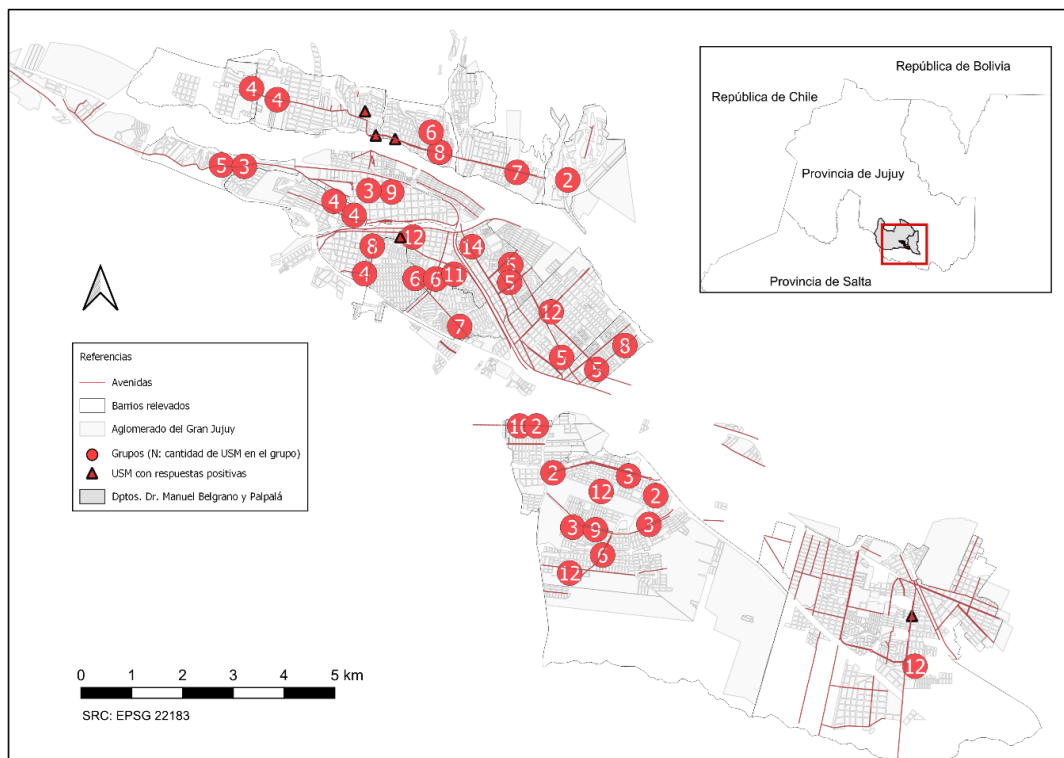


Figura 1: Distribución de USM por barrio

Los registros se tomaron con el aplicativo *KoboCollect*, de la plataforma *KoboToolBox*, que permite coleccionar datos georeferenciados *offline* y posteriormente almacenarlos en una base de datos en la nube.

Siguiendo la metodología descrita por Vicentín Masaro *et al.*, (2020), en cada relevamiento $t \in \{1, 2, \dots, 7\}$, se tomó como medida de dispersión al *logaritmo de la relación del precio de cada producto, sobre el promedio de precios de ese producto*:

$$dv_{ijmt} = \ln \left(\frac{p_{ijmt}}{\bar{p}_{it}} \right) \quad (1)$$

donde p_{ijmt} es el precio del i -ésimo producto ($i = \{1, 2, \dots, 7\}$), en el j -ésimo punto de venta ($j = \{1, 2, \dots, K\}$) localizado en el m -ésimo barrio ($m = \{1, 2, \dots, M\}$) durante el momento t ($t \in \{1, 2, \dots, 7\}$), y l es el tipo de comercio para su identificación ($l = \{1, 2, \dots, L\}$); y $\bar{p}_{it} = \frac{1}{MKL} \sum_{m=1}^M \sum_{j=1}^K \sum_{l=1}^L p_{ijmt}$ es el precio promedio del i -ésimo producto en el período t . Se puede notar que en (1) es equivalente a calcular dv_{ijmt} con precios deflactados y sin deflactar. En adelante, se indicará dv_{ijmt} simplemente como dv . Un punto de venta con un $dv = 0$ para un determinado producto en el tiempo t , indica que ese punto de venta presentó un precio igual al precio promedio para ese producto en el tiempo t .

La determinación de las contribuciones relativas de los factores *barrio* y *tipo de comercio* sobre el valor de dv condicional del producto i en el tiempo t , se realizó mediante la construcción de un *modelo de efectos aleatorios*, con la forma:

$$Y_j | i, t = \mu_{it} + M_{ijt} + L_{ijt} + \epsilon_{ijt} \quad (2)$$

Donde $Y_j | i, t$ representa el valor de dv del producto i en el tiempo t , M_{ijt} y L_{ijt} son los efectos aleatorios marginales de barrio y tipo de comercio respectivamente. Los modelos de efectos aleatorios

son denominados así en el sentido de que sus estimaciones son consideradas como resultados aleatorios o estocásticos del proceso que intenta estimar y no como constantes estimables (Gelman & Hill, 2006). Este enfoque permite aislar la variabilidad estructural de la variabilidad residual incluso en diseños no balanceados (Gelman & Hill, 2006). Esta estructura permite descomponer la varianza total de Y en 3 (tres) componentes: τ^2 , γ^2 y σ^2 , asignables al barrio, tipo de comercio y punto de venta, respectivamente. Cada varianza parcial mide la proporción de la variación total que puede ser explicada por cada componente. El ajuste del modelo se realizó con la función `lmer()` y el cálculo de las estimaciones relativas de los componentes de varianza se realizó con la función `VarCorr()`, ambas funciones de la librería {lme4} del software estadístico R (4.4.3).

A continuación, se realizó un *análisis de Procrustes Generalizado* (GPA, por sus siglas en inglés), que es una técnica estadística multivariada utilizada para comparar y sintetizar matrices de datos multidimensionales. Para ello se empleó la función `GPA` de la librería {FactoMineR} de R (4.4.3). Este método permite obtener una combinación lineal de los datos llamada *configuración de consenso* o promedio a partir de dos o más matrices provenientes de análisis de componentes principales (PCA), denotadas como, $X_1, X_2, \dots, X_t$, correspondientes a distintos momentos de observación t . La transformación Procrustes se obtiene por escalamiento, rotación y traslación de una matriz de datos que, en términos matriciales, pueden ser expresados del siguiente modo:

$$Y_k = \rho_k C_k H_k + T_k \quad (3)$$

donde Y_k representa la transformación de Procrustes, ρ_k el factor de escala, C_k es una matriz resultante de un Análisis de Componentes Principales (ACP) aplicado sobre una matriz X_t de datos conformadas por n filas (individuos) y p columnas (variables), H_k es la matriz ortogonal de rotación de dimensión $p \times p$ y T_k la matriz traslación de dimensión $n \times p$ (Commandeur, 1991; Bruno & Balzarini, 2010; Del Médico, & Vitelleschi, s. f.). En adelante, se designará a la matriz de consenso simplemente como *matriz Z*. Esta matriz puede ser interpretada como un conjunto de valores o *cargas* cuyas magnitudes y direcciones permiten asociar a las unidades observadas (en este trabajo, esas unidades son los barrios) según un criterio de distancias, de la forma que se realiza usualmente en un ACP. Como resultado del GPA, fue posible construir biplots a partir de la matriz Z , que permitieron reconocer patrones y asociaciones por proximidad espacial entre variables. También fue posible analizar la estructura de la variabilidad mediante la descomposición de la Suma de Cuadrados Total (SCT) por medio de un Análisis de la Varianza de Procrustes (PANAVA), que permite particionar la variación total observada en componentes atribuibles al consenso y a los residuos, esto es, los desvíos de cada mes respecto al consenso. Esta descomposición proporciona una medida de la “estabilidad” de las configuraciones espaciales a lo largo del tiempo (Gower, 1975).

El enfoque de Procrustes tiene la ventaja de que es posible trabajar con configuraciones parciales incompletas, es decir, con datos faltantes u omitidos (Commandeur, 1991; Gower, 2005), por lo que puede aplicarse a los datos obtenidos en el relevamiento realizado en este trabajo.

A fin de identificar agrupamientos o estratos de barrios con valores homogéneos de la variable de respuesta (dv), se realizó un análisis de clústeres jerárquico sobre la matriz Z . Se empleó el método de Ward, que asegura la mínima varianza dentro de cada clúster (Ward, 1963). Se construyeron *dendrogramas* para representar las relaciones entre los barrios por mayor proximidad de dv, sobre los cuales se identificaron aquellos agrupamientos que maximizaron la variabilidad *entre* agrupamientos y minimizaron la variabilidad *dentro* de cada agrupamiento. El número g de estos agrupamientos fue establecido mediante el método de la *silueta*, y resulta del número de agrupamientos que maximiza el *coeficiente de la silueta* para un punto i , definido como:

$$s(i) = \begin{cases} 1 - a(i)/b(i), & \text{si } a(i) < b(i) \\ 0, & \text{si } a(i) = b(i) \\ b(i)/a(i) - 1, & \text{si } a(i) > b(i) \end{cases} \quad (4)$$

donde $a(i)$ es la distancia del punto i respecto de todos los puntos del mismo agrupamiento, y $b(i)$ es la distancia del punto i y los puntos del grupo más cercano al que no pertenece (Rousseeuw, 1987). Se debe notar que $s(i)$ varía entre -1 y 1.

Una vez identificados los agrupamientos de barrios por dispersión de precios similares, se calcularon el *Índice de Moran Global* y el *Índice de Moran Local*, también llamado *Local Indicators of Spatial Association* (LISA), que es una medida para identificar patrones de autocorrelación espacial a nivel *local* ya que, a diferencia del Índice de Moran Global (que evalúa la autocorrelación espacial en todo el conjunto de datos), LISA detecta clústeres espaciales (áreas con valores similares) y outliers espaciales (valores atípicos en relación con sus vecinos). Se utilizaron las funciones `moran.test()` y `localmoran()` de la librería {spdep} de R (4.4.3).

El primero se calcula:

$$I = \frac{n \sum_{(2)} w_{ij} z_i z_j}{S_0 \sum_{i=1}^n z_i^2}$$

En tanto, el Índice de Moran Local se calcula:

$$I_i = \frac{N}{\sum_{i=1}^N (x_i - \bar{x})^2} \frac{x_i - \bar{x}}{m_2} \sum_{j=1}^N w_{ij} (x_j - \bar{x})$$

donde w_{ij} es una matriz de ponderaciones espaciales, $z_j = (x_j - \bar{x})$ y $m_2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$ (Pebesma & Bivand, 2023). La matriz w_{ij} fue calculada con el criterio *Queen* ("reina"), según el cual dos áreas o *Láttices* pueden ser considerados vecinos si comparten al menos un borde o vértice. Luego, sus centroides fueron utilizados para definir la matriz de pesos espaciales, asignando un valor positivo solo a aquellas unidades que cumplen con dicha relación de vecindad.

En función del valor de I_i , una unidad o *Láttice* puede ser clasificado según su relación con sus vecinos según las categorías indicadas en Tabla 4:

Tabla 4: Categorías de *Láttices* según el resultado de LISA

Categoría	Signo de I_i	Signo de I_j (vecinos)	Interpretación
Alta - Alta (HH)	Positivo	Positivo	Área con valor alto rodeada de valores altos (clúster de altas)
Baja - Baja (LL)	Negativo	Negativo	Área con valor bajo rodeada de valores bajos (clúster de bajas)

Alta - Baja (HL)	Positivo	Negativo	Área con valor alto rodeada de valores bajos (outlier espacial)
Baja - Alta (LH)	Negativo	Positivo	Área con valor bajo rodeada de valores altos (outlier espacial)

Fuente: Moraga, (2024)

RESULTADOS Y DISCUSIÓN

En la Figura 2 se observan los valores tipificados de *dv* por barrio, según producto analizado. En términos generales, desde un enfoque univariado, la dispersión de los precios por barrio se distribuyó en torno a los valores -2 y 2 desvíos estándar, y el 50% central de los precios se ubicó entre -1 y 1 desvío estándar del valor medio, lo que sugiere un comportamiento general de tipo normal o gaussiano de la dispersión de los precios en torno al valor medio o de equilibrio. Dentro de este comportamiento normal, algunos barrios tendieron a mostrar un *dv* por encima del valor medio (Los Perales, Alto y Bajo La Viña, Ciudad de Nieva, Dr. Marcelino Vargas y Centro). Los valores de *dv* registrados en el barrio Los Perales (L.P.) para lechuga se ubicaron consistentemente por encima del percentil 97.5, sugiriendo un comportamiento estadísticamente atípico.

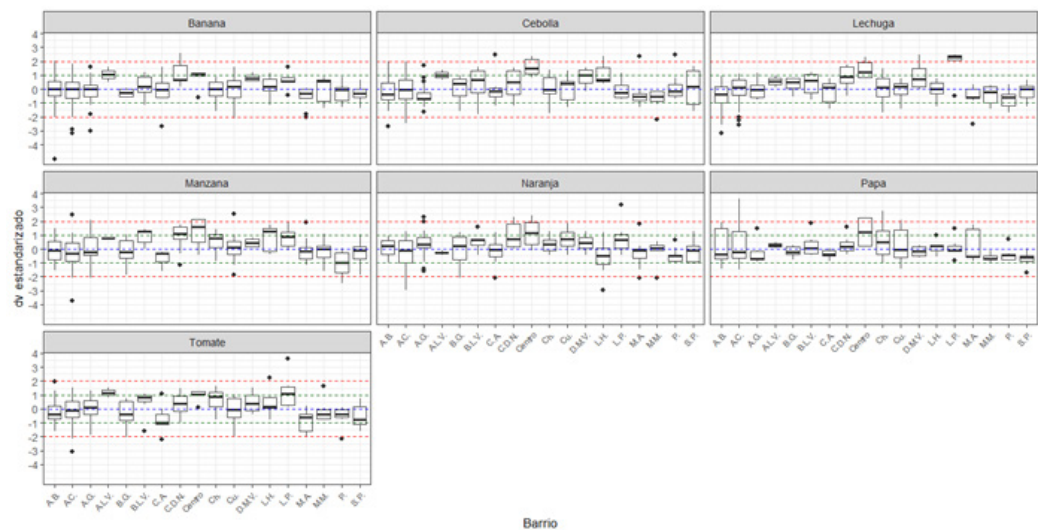


Figura 2: Dispersión de precios (*dv*) estandarizada por barrio, según producto

Para establecer la contribución relativa de los factores barrio y tipo de comercio sobre *dv*, se realizó una descomposición de los componentes de varianza (Gelman & Hill, 2006; Snijders & Bosker, 2012) utilizando un modelo lineal de efectos aleatorios. La bondad de ajuste fue evaluada mediante el diagnóstico de los residuos estandarizados; su distribución es compatible con el supuesto de normalidad (Pinheiro & Bates, 2000), lo que permite determinar la validez del modelo para la inferencia estadística, tal como se verifica en la Figura 3:

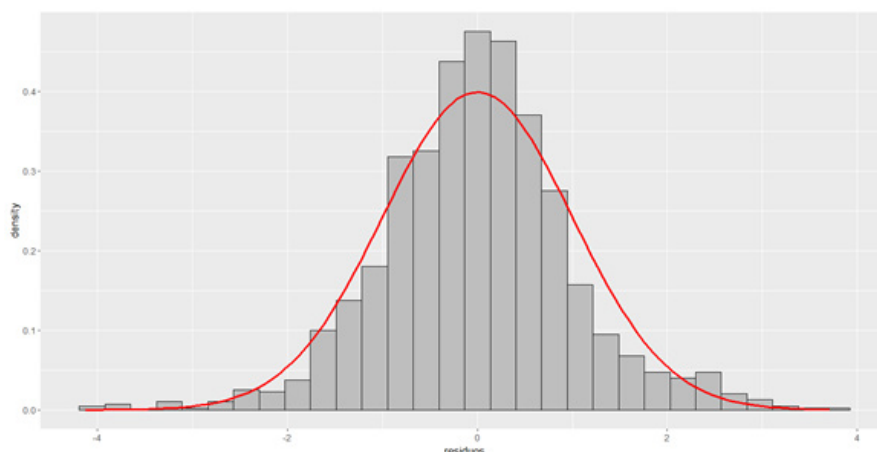


Figura 3: Distribución empírica y teórica normal de los residuos estandarizados del modelo de efectos aleatorios ajustado

A continuación, se presentan las contribuciones a la varianza de cada factor a nivel general, y por producto (Tabla 5):

Tabla 5: Contribuciones relativas de las componentes de varianza asociadas a cada factor

Producto	Barrio σ^2	Tipo de Comercio σ^2	Residual (punto de venta) σ^2
	%		
Papa	10,15	0,29	89,56
Banana	16,02	3,49	80,49
Cebolla	9,24	2,23	88,53
Lechuga	31,22	1,58	67,2
Naranja	9,97	0	90,03
Manzana	19,45	6,79	73,76
Tomate	16,64	15,88	67,48
General	14,94	1,17	83,89

En términos generales, en un momento de relevamiento *t* dado, la variabilidad explicada por el factor barrio representó el 15% de la variación total observada, mientras que el factor tipo de comercio tuvo una contribución muy baja, de poco más del 1%. La variabilidad general no explicada por el modelo fue de aproximadamente un 84%, la cual puede ser imputada al punto de venta como un término que sintetiza una sumatoria de efectos de variables no incluidas en el modelo y que se puede interpretar como una variabilidad intra barrio. Lo anterior pone en perspectiva la contribución moderada del factor barrio como variable explicativa en la formación de precios minoristas, y prácticamente nula del tipo de comercio (nula en el caso de la naranja), con excepción del tomate. La papa, cebolla y naranja fueron los productos relativamente más estables entre barrios, mientras que la lechuga fue la más variable entre barrios, seguida por la manzana (Tabla 5).

Este resultado constituye una evidencia empírica de la alta dispersión intra barrial, con fuertes asimetrías en la conformación de precios minoristas. Estas asimetrías, potenciadas por ausencia o escasez de información sobre precios de referencia, impiden la convergencia hacia un precio único,

incluso dentro de un mismo distrito o barrio. No obstante, aunque moderada, la contribución relativa del factor barrio puede considerarse como una fuente de variación estructural o componente espacial con efectos relativamente estables en el costo de acceso a los productos analizados en el aglomerado.

Dado que el factor tipo de comercio tuvo una contribución relativa muy baja para explicar la variabilidad general observada, se realizó el análisis GPA considerando únicamente las dimensiones barrio y momento t. La Tabla 6 detalla las componentes de las sumas de cuadrados marginales por barrio de un PANAVA en la que pueden observarse las componentes de varianza entre y dentro de cada momento de relevamiento t, para cada barrio:

Tabla 6: Descomposición de sumas de cuadrados de Procrustes marginales por barrio

Barrio	SC Ajuste	SCR	SCT	%
Centro	0,90	0,10	1,00	89,63
P.	0,32	0,04	0,36	88,27
L.P.	0,62	0,12	0,74	84,18
A.L.V.	0,07	0,01	0,08	83,83
D.M.V.	0,17	0,05	0,23	76,39
M.M.	0,26	0,10	0,36	70,94
B.L.V.	0,30	0,16	0,46	65,73
M.A.	0,21	0,11	0,32	65,11
A.B.	0,30	0,16	0,46	64,84
C.D.N.	0,31	0,18	0,49	62,49
L.H.	0,37	0,24	0,61	60,94
B.G.	0,06	0,04	0,10	58,51
A.G.	0,21	0,16	0,37	56,27
C.A.	0,26	0,22	0,47	54,18
S.P.	0,20	0,20	0,40	50,63
Cu.	0,05	0,07	0,12	43,04
A.C.	0,13	0,18	0,31	41,21
Ch.	0,08	0,18	0,26	29,30
sum	4,80	2,33	7,13	67,29

La columna ‘%’ de la Tabla 6, ordenada en forma descendente, es una medida de la “estabilidad” relativa entre momentos de muestreo de la dispersión de precios por barrio. En este estudio, el barrio Centro mostró la menor variabilidad en los precios a lo largo del período de relevamiento, lo que indica un comportamiento más estable. En contraste, el barrio Chijra presentó la mayor dispersión de precios entre relevamientos, lo que sugiere un comportamiento más inestable.

Desde una perspectiva metodológica, la estabilidad relativa entre momentos es un indicador de la representatividad de los componentes de la matriz Z para describir la variabilidad subyacente a cada barrio (a mayor estabilidad relativa, mayor representatividad), y, en consecuencia, de la confiabilidad de las conclusiones que a partir de ella se deriven.

Desde el punto de vista del análisis de precios, esta métrica permite distinguir entre variabilidad estructural y volatilidad coyuntural. Una mayor estabilidad relativa sugiere la existencia de mecanismos de formación de precios con un mayor componente determinista; por el contrario, una baja estabilidad sugiere una formación de precios errática, donde la dispersión no responde a un patrones regulares, sino a ineficiencias por falta de información o a la informalidad propia de esos puntos de venta.

Como resultado de esta modelización, se obtuvo la matriz Z que, como se indicó anteriormente, permite analizar proximidades entre barrios en términos de distancias, para un momento tipo t , de forma análoga a un Análisis de Componentes Principales (ACP). A fin de establecer el número óptimo de agrupamientos homogéneos entre barrios, se realizó la prueba de la silueta, que permitió identificar 2 (dos) estratos o agrupamientos homogéneos (Figura 4):

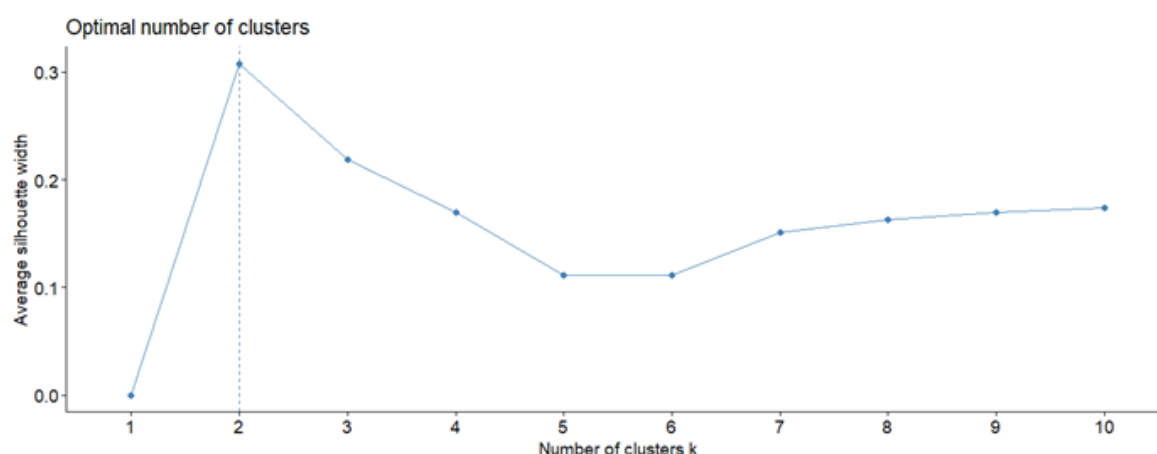


Figura 4: Resultado de la prueba de la silueta, sobre la matriz Z

Tabla 7: Variabilidad explicada por cada dimensión de la matriz Z

Variación	DIM1	DIM2	DIM3	DIM4	DIM5	DIM6	DIM7
Desviación estándar	0,17	0,10	0,08	0,06	0,06	0,05	0,04
Proporción de varianza	0,51	0,19	0,10	0,07	0,06	0,04	0,03
Proporción acumulada	0,51	0,70	0,80	0,87	0,93	0,97	1,00

En Tabla 7 se observa que las dos primeras dimensiones de la matriz Z explican un 70% de la variación total observada, lo cual es un buen valor descriptivo, ya que supera el umbral del 60% sugerido habitualmente en las ciencias sociales para considerar una solución factorial como satisfactoria (Hair et al., 2019). En consecuencia, se construyó un gráfico *biplot* (Figura 5) con las dos primeras componentes de Z, identificando los barrios de cada uno de los dos clústeres:

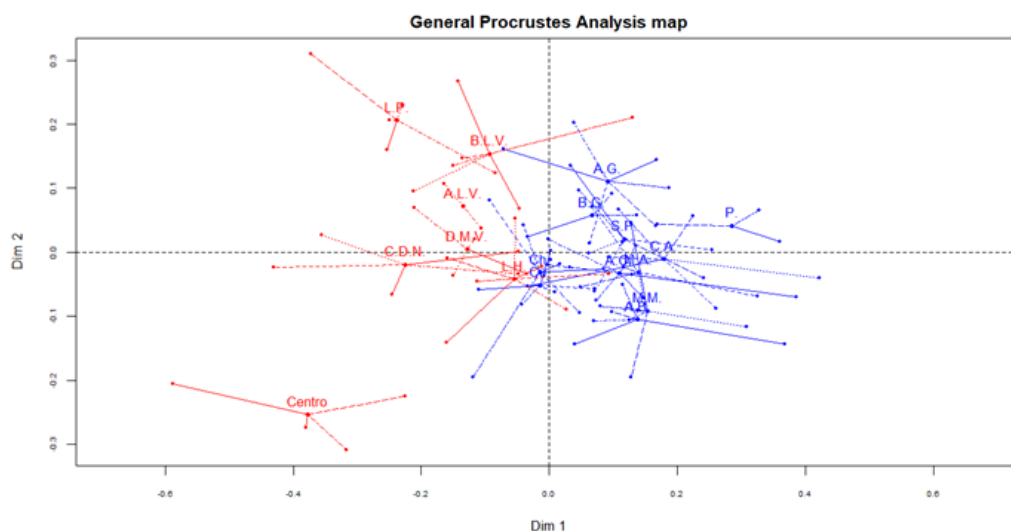


Figura 5: Biplot de configuraciones parciales y de consenso de la matriz Z

El biplot construido a partir de la matriz Z permite visualizar la relación entre las configuraciones parciales de cada barrio y la configuración de consenso (Figura 5). Esta relación se presenta como una “estrella”, donde los nodos satélite corresponden a las configuraciones parciales obtenidas en cada momento t de relevamiento, y el nodo central corresponde al promedio de las anteriores, o consenso. Cuanto más cerca entre sí los nodos centrales, mayor es la proximidad o semejanza en la dispersión de precios entre barrios. Sin embargo, se hace dificultoso poder identificar y comparar visualmente relaciones de proximidad entre barrios. Por ese motivo, se realizó un análisis de *clústeres jerárquicos* por el método de Ward, que permitió construir un Dendrograma adecuado para facilitar la observación de estas proximidades.

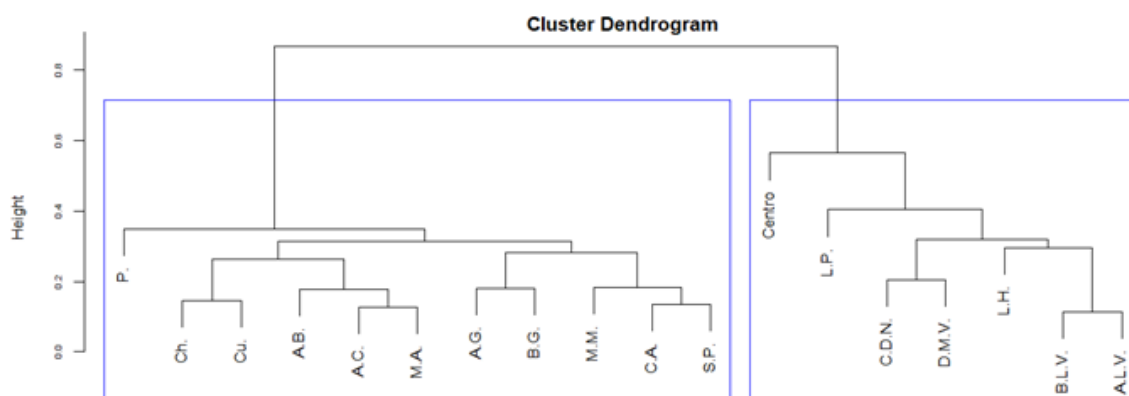


Figura 6: Dendrograma de un análisis clúster jerárquico por el método de Ward, sobre las cargas de cada barrio en la matriz Z

La Figura 6 permite identificar con mayor claridad las asociaciones entre barrios según la dispersión general de precios. La segmentación en dos estratos denota un comportamiento diferencial de los barrios del distrito Norte, junto con Los Huaicos y San Salvador de Velazco en el Valle de Jujuy. A su vez, dentro de cada clúster, pueden reconocerse cierta estructura subyacente de relación entre barrios contiguos. A fin de determinar si estas relaciones son significativas, se realizó la prueba de Moran Global con una matriz de pesos “Queen” (Figura 7), obteniéndose un valor $p = 0,003$, que resultó *altamente significativo* con un valor $p = 0,003$ (Cliff & Ord, 1981). Por lo tanto, *existe una autocorrelación espacial positiva global*, lo que indica que los barrios con precios similares tienden a agruparse geográficamente:

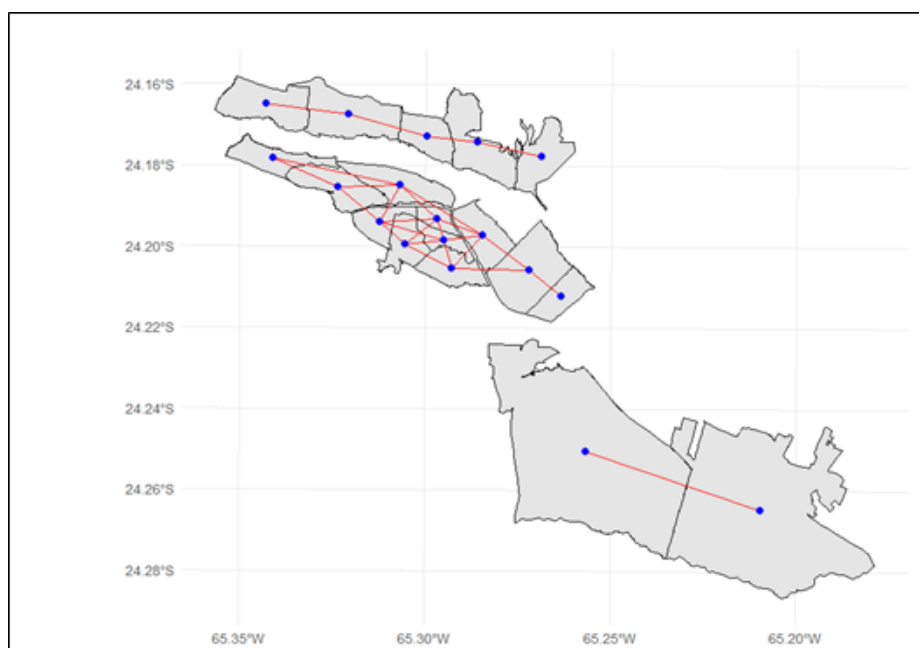


Figura 7: Red de centroides conectados por vecindad según el criterio "Queen"

Se realizó la prueba LISA o *Prueba Local de Moran* sobre la primera dimensión de la matriz Z que, como se vio anteriormente permite explicar un 51% de la variabilidad observada y constituye el eje sobre el cual puede identificarse la "segmentación" de los clústeres. El resultado indica que fueron significativas las autocorrelaciones locales en Los Huaicos y Ciudad de Nieva, con valores de I_{local} de 0,59 ($p = 0,01$) y 1,13 ($p = 0,04$) respectivamente. En ambos casos se trató de correlaciones *High-High*, es decir, barrios con dv elevada rodeados por barrios con dv elevados.

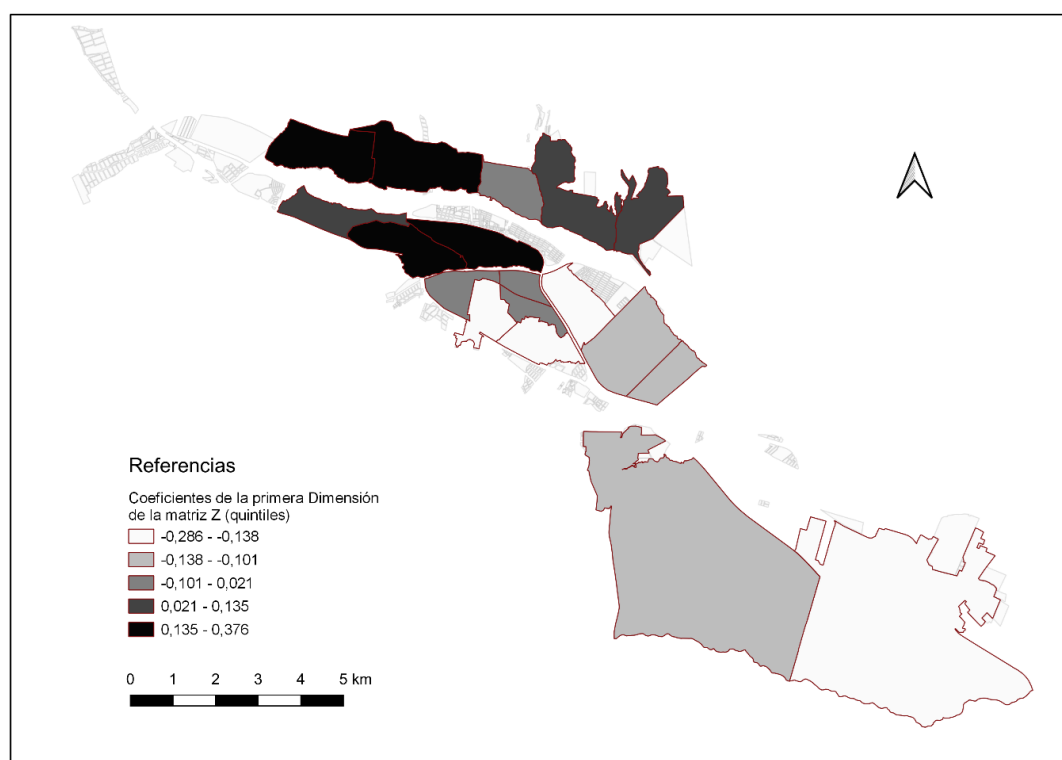


Figura 8: Quintiles de coeficientes o *loadings* de la primera dimensión de la matriz Z por barrio

En la Figura 8 se observan, en tonos oscuros, los barrios con coeficientes de carga o *loadings* mayores en la primera dimensión de la matriz Z, asociados con precios relativos más elevados. En tonos claros, se identifican los barrios con precios relativos menores.

Esta configuración territorial permite evidenciar diferencias estructurales con incidencia en los costos de comercialización y la formalidad de canales de venta. Los barrios ubicados hacia el sur del aglomerado, asociados a menores precios relativos, coinciden con zonas donde prevalecen circuitos de comercialización más informales como ferias barriales o venta ambulante. Por el contrario, los precios altos en los barrios del centro y norte del aglomerado, pueden atribuirse, entre otros factores, a mayores costos fijos de comercialización (alquileres comerciales) y a una demanda menos elástica al precio, lo que permite a los oferentes sostener precios superiores de manera sistemática.

Finalmente, los resultados de este análisis exploratorio representan un punto de partida para continuar esta línea de investigación, integrando otras variables con potencial poder predictivo y explicativo, a fin de comprender en mayor profundidad la variabilidad en la dispersión de precios minoristas de frutas y verduras en el aglomerado Jujuy – Palpalá.

CONCLUSIONES

Los resultados obtenidos en este estudio demuestran que, incluso en un contexto caracterizado por la informalidad y la escasez de información en la cadena de valor hortícola provincial, fue posible identificar patrones espaciales estables en la formación de precios minoristas. La evidencia analizada sustenta la hipótesis de que el factor barrio ejerce un peso relativo moderado pero estable como variable explicativa de la variabilidad de precios. Por otra parte, el tipo de comercio no presentó aportes significativos, con la excepción de productos específicos como el tomate.

El análisis multivariado aplicado sobre la matriz de consenso evidenció la existencia de dos agrupamientos espaciales bien definidos, que separan geográficamente al aglomerado Jujuy-Palpalá, destacándose un patrón diferencial en los barrios del distrito Norte, junto con Los Huaicos y barrio Centro. Esta segmentación fue validada mediante la prueba de autocorrelación espacial global de Moran I, cuyo valor significativo confirma que los precios no se distribuyen aleatoriamente en el territorio, sino que tienden a agruparse espacialmente. Simultáneamente, la prueba LISA permitió detectar clústeres locales relevantes en Ciudad de Nieva y Los Huaicos, corroborando la existencia de núcleos de precios relativos mayores rodeados de zonas con comportamientos similares.

En línea con los argumentos desarrollados en la introducción, donde se plantea que la dispersión de precios es una manifestación de información imperfecta en mercados atomizados e informales, los hallazgos de este estudio refuerzan la necesidad de contar con sistemas de monitoreo georreferenciados y metodologías analíticas robustas que aporten a una toma de decisiones más eficiente. Este enfoque aporta evidencia para comprender mejor el funcionamiento del mercado minorista de productos para consumo fresco, lo que puede constituir una herramienta de utilidad para la toma de decisiones informadas.

BIBLIOGRAFÍA

Bruno, C. & Balzarini, M. (2010). Ordenaciones de material genético a partir de información multidimensional. Revista de la Facultad de Ciencias Agrarias. Universidad Nacional de Cuyo. 42(2): 183-200.

Cliff, A. D., Ord, J. K. 1981 Spatial processes, Pion, p. 21.

- Commandeur, J. (1991). *Matching Configurations*. Department of Psychometrics and Research Methodology. DSWO Press, Leiden University. Países Bajos.
- Del Médico, Ana Paula & Vitelleschi, María Susana. (s. f.). *Análisis procrustes generalizado. Una aplicación en el Área agrícola*. Recuperado de: <https://rephip.unr.edu.ar/items/65620950-5641-4adc-b090-35da0ae2f1e8/full>
- Dirección General de Programas y Proyectos Sectoriales y Especiales - MAGYP. (2022). *EPSA - ESTRATEGIA PROVINCIAL PARA EL SECTOR AGROALIMENTARIO*. Recuperado de:
https://www.argentina.gob.ar/sites/default/files/2022/08/02_epsa_jujuy.pdf
- Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models* (1.^a ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511790942>
- Gower, J. y Dijksterhuis, G. 2005. *Procrustes problems*. Oxford University Press.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.): 142. Cengage Learning.
- Instituto Nacional de Estadística y Censos (Argentina) (2020). *Precenso de viviendas*. Recuperado de <https://precensodeviviendas.indec.gob.ar/>
- Instituto Nacional de Estadística y Censos (Argentina) (Ed.). (2021). *Censo Nacional Agropecuario 2018: Resultados definitivos, abril de 2021*. Ministerio de Economía, Argentina : Instituto Nacional de Estadística y Censos, República Argentina. Recuperado de:
https://www.indec.gob.ar/ftp/cuadros/economia/cna2018_resultados_definitivos.pdf
- Market Information Organization of the Americas, Washington, D.C. (United States of America) & Instituto Interamericano de Cooperación para la Agricultura (IICA). (2016). *Manual sobre análisis básico de precios agrícolas para la toma de decisiones*.
- Masaro, J. V., Rossini, G., & Chara, A. L. (2020). EVIDENCIA DE LA DISPERSIÓN DE PRECIOS DE ALIMENTOS MINORISTA EN SANTA FE - ARGENTINA. *Revista Americana de Emprendedorismo e Inovação*, 2(1), 156-165. <https://doi.org/10.33871/26747170.2020.2.1.3286>
- Moraga, P. (2024). *Spatial statistics for data science: Theory and practice with R* (First edition). CRC Press.
- Organización de las Naciones Unidas para la Alimentación y la Agricultura. (2023). *América Latina y el Caribe—Panorama Regional de la Seguridad Alimentaria y la Nutrición 2023*. FAO; IFAD; PAHO; WFP; UNICEF; <https://doi.org/10.4060/cc8514es>
- Pebesma, E., & Bivand, R. (2023). *Spatial Data Science: With Applications in R* (1.^a ed.). Chapman and Hall/ CRC. <https://doi.org/10.1201/9780429459016>
- Pérez, G.A. (2010). *El mercado minorista de hortalizas en la provincia de Tucumán* [Universidad de Buenos Aires]. Recuperado de: <https://repositorio.inta.gob.ar/xmlui/handle/20.500.12123/6754>

- Pinheiro, J. C., & Bates, D. M. (2000). Mixed-effects models in S and S-PLUS. Springer. <https://doi.org/10.1007/b98882>
- Martínez, R., Golovanevsky, L. & Medina, F.. (2010). *Economía y empleo en Jujuy*. Recuperado de: <https://repositorio.cepal.org/server/api/core/bitstreams/68811698-60a4-4e60-bd42-d6210304bd4a/content>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Snijders, T. A. B., & Bosker, R. J. (2012). Multilevel analysis: An introduction to basic and advanced multilevel modeling (2.ª ed.). SAGE Publications.
- Sorensen, A. T. (2000). Equilibrium Price Dispersion in Retail Markets for Prescription Drugs. *Journal of Political Economy*, 108(4), 833-850. <https://doi.org/10.1086/316103>
- Stahl, D. O. (1996). Oligopolistic pricing with heterogeneous consumer search. *International Journal of Industrial Organization*, 14(2), 243-268. [https://doi.org/10.1016/0167-7187\(94\)00474-9](https://doi.org/10.1016/0167-7187(94)00474-9)
- Stigler, G. J. (1961). The Economics of Information. *Journal of Political Economy*, 69(3), 213-225. <https://doi.org/10.1086/258464>
- Varian, H. (1981). A Model of Sales: Errata. *American Economic Review*, 71(3). <https://EconPapers.repec.org/RePEc:aea:aecrev:v:71:y:1981:i:3:p:517>
- Vicentin Masaro, J. & Chara, A.L. (2020). ¿Existen patrones espaciales en la fijación de precios minoristas de alimentos en los supermercados de la ciudad de Santa Fe? *Cuadernos del Cimbage*, 22, 105-143.
- Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), 236-244. <https://doi.org/10.1080/01621459.1963.10500845>