

Construcción y validación del modelo de innovación

La transformación del sector cafetero colombiano hacia modelos más sostenibles y competitivos requiere herramientas que orienten eficazmente los procesos de innovación. En este capítulo se presenta el desarrollo sistemático de un modelo de innovación diseñado específicamente para el contexto de las fincas cafeteras productoras de café especial e

n Colombia.

Partiendo de la caracterización realizada en el capítulo anterior, se avanza hacia la construcción de un modelo que no solo identifica las variables críticas que determinan la capacidad innovadora, sino que permite predecir, con alta precisión la probabilidad de que una finca pueda transformarse en productora de café especial. Este modelo representa una herramienta práctica para caficultores, técnicos, cooperativas y entidades de fomento interesadas en promover la innovación como vía para mejorar la calidad, sostenibilidad y rentabilidad de la producción cafetera.

6.1 Fundamentos metodológicos del modelo

6.1.1 Enfoque conceptual

El desarrollo del modelo de innovación se fundamenta en una combinación de enfoques teóricos y prácticos que buscan capturar la complejidad

multidimensional de la innovación en el contexto cafetero. Como base conceptual, se adoptan los tipos de innovación propuestos en el Manual de Oslo (2018) los cuales se definen como la implementación de un producto o servicio nuevo o significativamente mejorado, cambios en los procesos productivos o innovaciones en las formas de de relacionamiento.

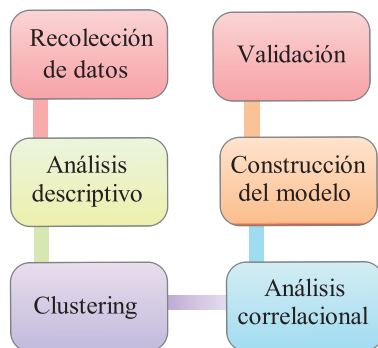
Esta conceptualización se complementa con el enfoque de Lambrecht et al., (2014), que plantea una visión integral de la innovación, incorporando tanto aspectos tecnológicos como no tecnológicos, considerando variables aplicables específicamente al contexto agrícola. Además, se incorporan elementos de los modelos de innovación sostenible propuestos por Samper y Quiñones (2017) quienes destacan que las certificaciones pueden contribuir al fortalecimiento de la innovación en el sector cafetero.

El estudio desarrollado identifica siete factores clave de innovación, relacionados con factores económicos (F1), sociales (F2), ambientales (F3), producción (F4), conocimiento (F5), tecnología (F6) y gestión del cambio (F7). Estos factores permiten comprender el relacionamiento entre variables que inciden en la capacidad de innovación en el proceso de un producto.

6.1.2 Diseño metodológico

El desarrollo del modelo siguió un proceso metodológico riguroso que combinó análisis cuantitativo y validación estadística. La figura 15 ilustra las etapas clave del proceso de construcción del modelo.

Figura 15. Etapas del proceso metodológico para la construcción del modelo



Nota. Elaboración propia.

El diseño metodológico incluyó las siguientes fases:

1. **Recolección de datos:** aplicación del instrumento validado a 66 fincas cafeteras seleccionadas mediante muestreo aleatorio simple.
2. **Análisis descriptivo:** caracterización de las fincas según su capacidad innovadora y producción de café especial.
3. **Clustering:** agrupación de fincas según su nivel de innovación para identificar patrones comunes.
4. **Análisis correlacional:** identificación de relaciones significativas entre variables de innovación y capacidad para producir café especial.
5. **Selección del modelo estadístico:** evaluación de diferentes aproximaciones para seleccionar el método más adecuado según la naturaleza de los datos.
6. **Construcción del modelo:** desarrollo del modelo predictivo basado en las variables más significativas.
7. **Validación:** verificación de la precisión y robustez del modelo mediante técnicas estadísticas.

Este enfoque secuencial y sistemático permitió garantizar la construcción de un modelo con sólido respaldo metodológico y capacidad predictiva verificable.

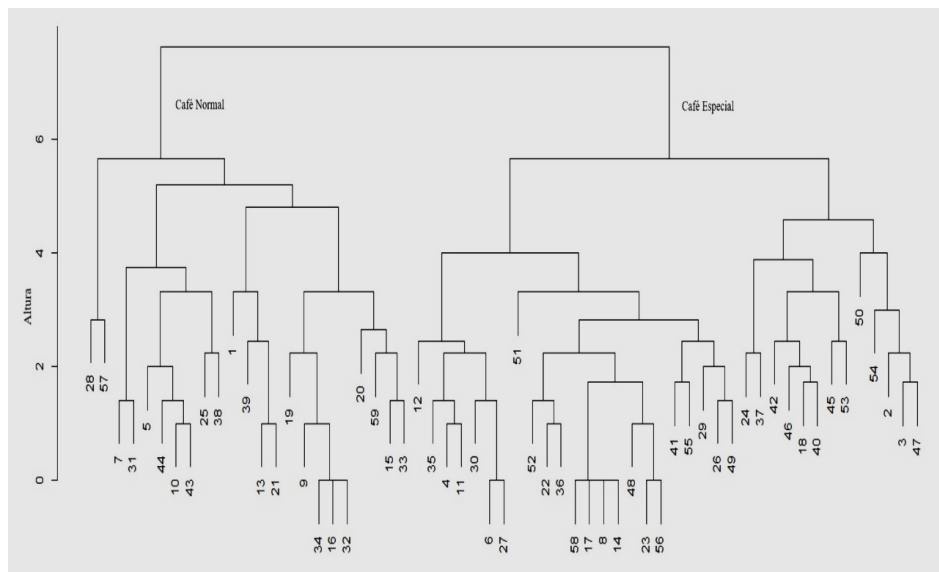
6.2 Análisis de agrupamiento e identificación de patrones

6.2.1 Técnicas de clustering aplicadas

El análisis de agrupamiento constituyó un paso fundamental para identificar patrones significativos entre las fincas estudiadas. Se desarrolló un análisis de agrupación jerárquica, utilizando distancia euclidiana como medida de similitud entre observaciones. Este enfoque permite formar grupos de fincas con características similares, a partir de sus variables de innovación.

La figura 16 muestra el dendrograma resultante que representa visualmente la estructura jerárquica de las fincas cafeteras.

Figura 16. Dendrograma de agrupación jerárquica de fincas que producen café convencional y especial



Nota. Elaboración propia con base en el procesamiento de datos mediante el software R-Commander.

El análisis de agrupamiento reveló la formación de dos clústeres principales:

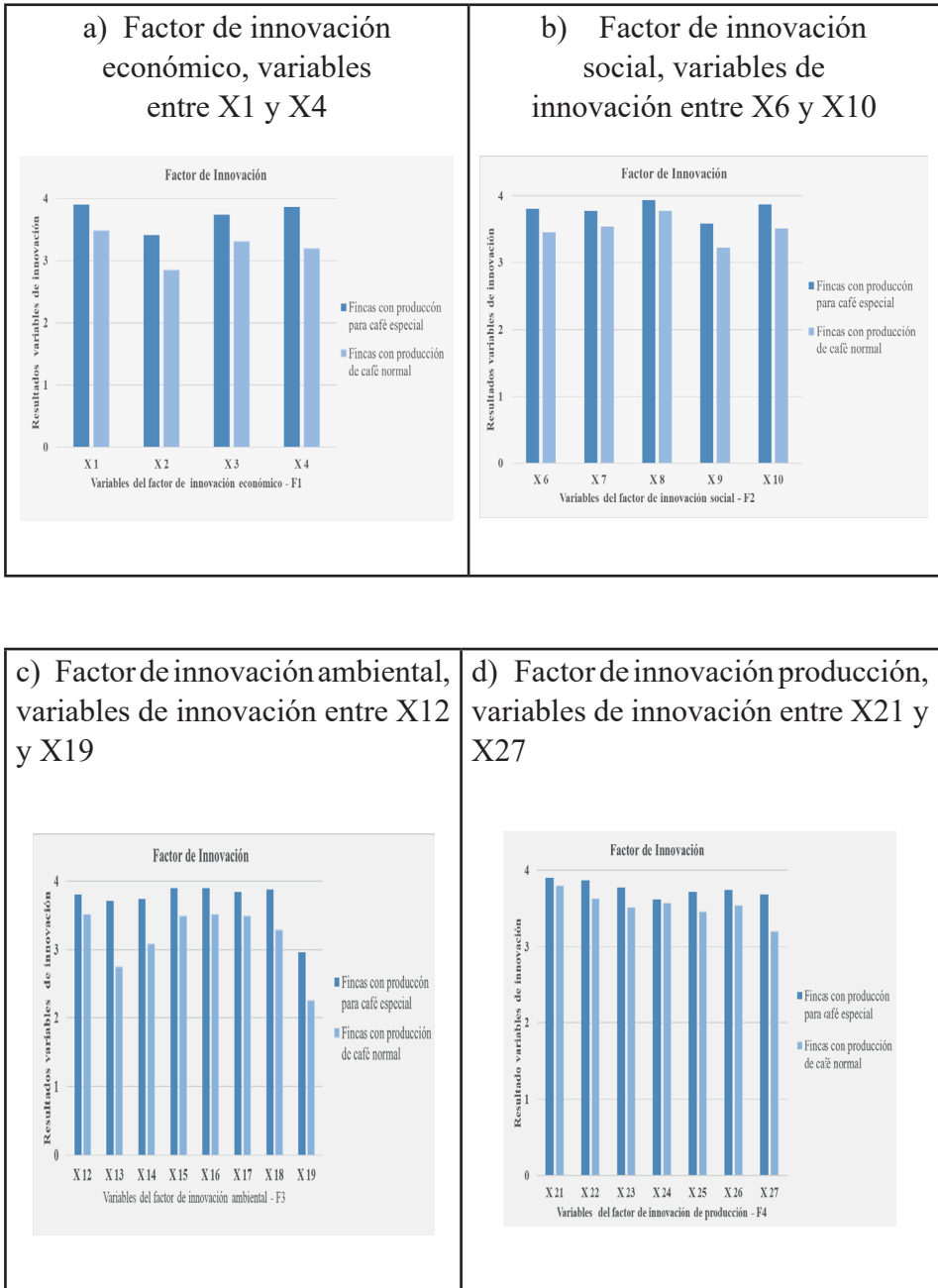
- **Clúster 1 - Fincas innovadoras:** conformado por 31 fincas (47 % de la muestra) caracterizadas por altos niveles en variables de innovación.
- **Clúster 2 - Fincas tradicionales:** integrado por 35 fincas (53 % de la muestra) con menores niveles de adopción de prácticas innovadoras.

Esta clasificación binaria se utilizó como variable dependiente para el desarrollo del modelo predictivo, etiquetando las fincas como innovadoras (valor 1) o no innovadoras (valor 0).

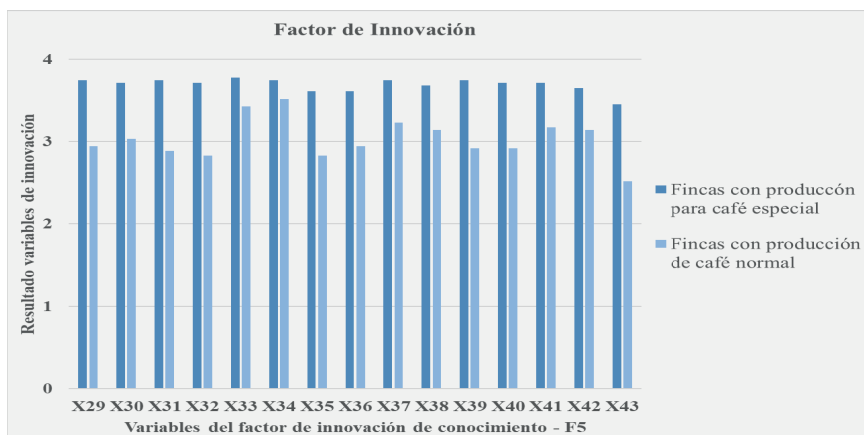
6.2.2 Diferenciación de perfiles entre grupos

El análisis comparativo entre los clústeres identificados evidenció diferencias significativas en los perfiles de innovación. La figura 17 representa la caracterización de los grupos según los siete factores de innovación analizados, permitiendo visualizar de manera comparativa los niveles de desempeño de cada grupo.

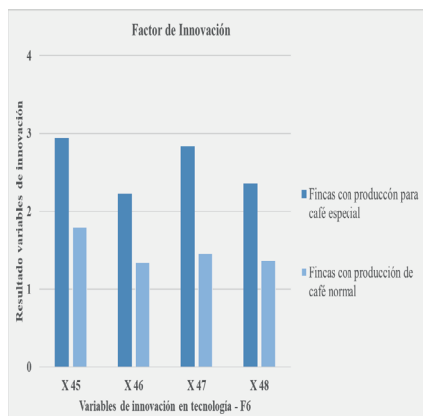
Figura 17. *Perfiles de innovación que tienen las fincas que producen café normal y especial*



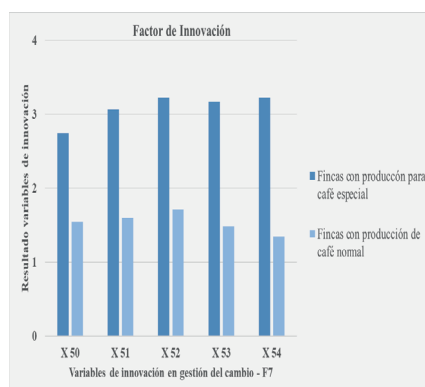
e) Factor de innovación conocimiento, variables de innovación entre X29 y X43



f) Factor de innovación en tecnología, variables de innovación entre X45 y X48



g) Factor de innovación en gestión del cambio, Variables de innovación entre X50 y X54



Nota. Elaboración propia.

Las mayores diferencias entre grupos se observaron en los factores de tecnología (F6) y gestión del cambio (F7), con brechas de 1.1 y 1.6 puntos respectivamente en una escala de 1 a 4. Estas diferencias marcadas sugieren que dichas dimensiones resultan particularmente determinantes en la clasificación

de una finca como innovadora. Al analizar las variables específicas dentro de cada factor, se identificaron tres variables que presentaron las mayores diferencias entre grupos:

- **X18 - Gestión de residuos** (Factor ambiental): Diferencia de 1.5 puntos.
- **X47 - Adopción digital** (Factor tecnológico): Diferencia de 1.8 puntos.
- **X51 - Flexibilidad al cambio** (Factor gestión del cambio): Diferencia de 1.9 puntos.

Estas variables emergieron como elementos diferenciadores clave entre los grupos, sugiriendo su potencial como predictores significativos para el modelo.

6.3 Análisis correlacional de variables de innovación

6.3.1 Métodos de correlación aplicados

Se realizó un análisis correlacional con el propósito de identificar las variables con mayor poder predictivo y comprender las relaciones entre las 54 variables de innovación y la probabilidad de producir café especial. Este análisis constituye un paso fundamental en la construcción del modelo de innovación dado que permite:

- Identificar variables con alta asociación con la variable dependiente.
- Detectar relaciones entre variables independientes (multicolinealidad).
- Validar los hallazgos del análisis ANOVA desde una perspectiva complementaria.
- Priorizar variables para la construcción del modelo predictivo.

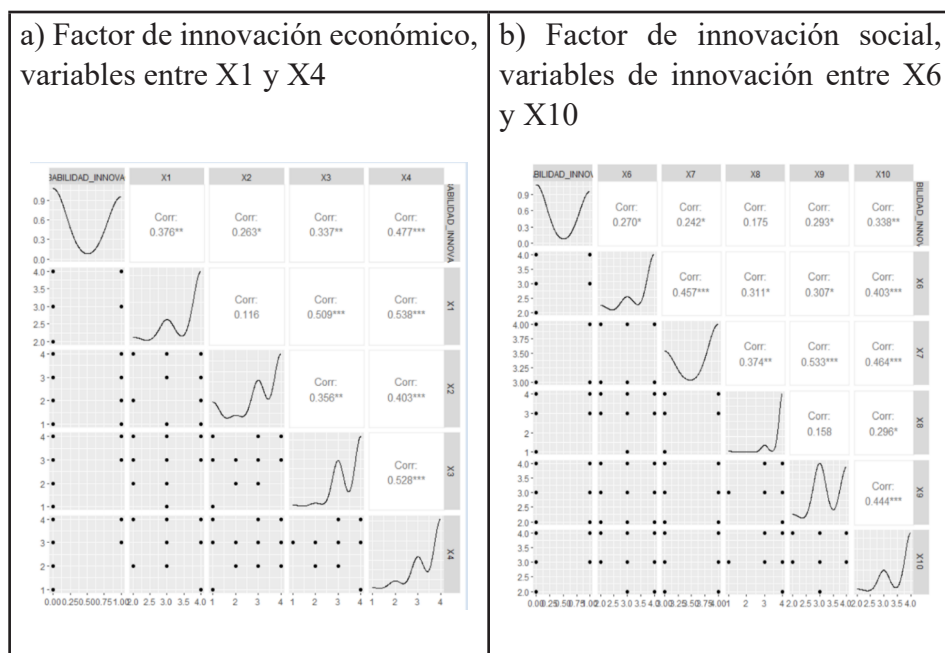
Para este análisis se aplicaron dos técnicas complementarias de correlación:

- **Análisis de emparejamiento (*Pairs*):** Esta técnica permite observar el relacionamiento de cada factor (variable independiente) con la variable probabilidad de innovación (Variable dependiente).

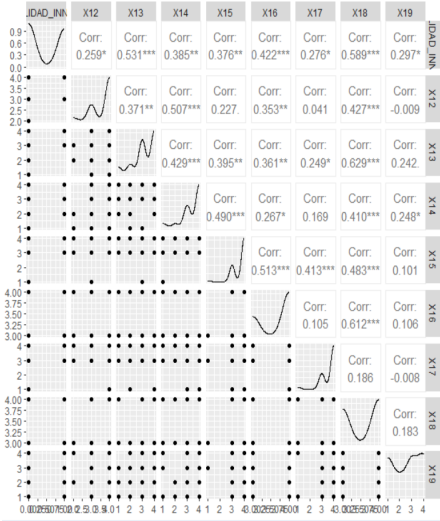
- **Mapas de calor (Heatmap):** Proporcionaron una representación visual de la matriz de correlación completa. que hacen más fácil detectar patrones y conexiones por medio de colores.

La figura 18 muestra gráficos de emparejamiento, evidenciando que la variable de probabilidad de Innovación (Y) muestra valores significativos de correlación con variables asociadas a os factores económico, ambiental, conocimiento, tecnología y gestión del cambio. Estas relaciones sugieren que dichas variables pueden incidir de manera directa en la mejora de los procesos productivos de las fincas cafeteras para generar un café especial.

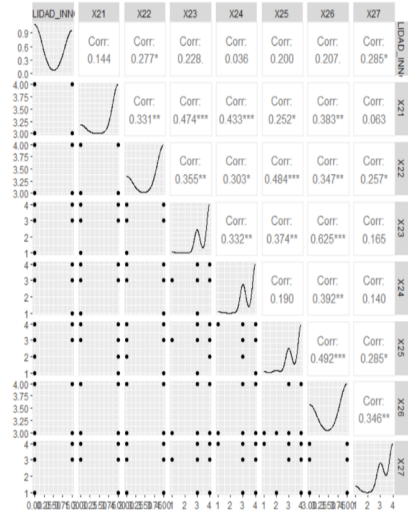
Figura 18. Análisis de emparejamiento entre probabilidad de innovación y variables que están relacionadas con los factores de innovación



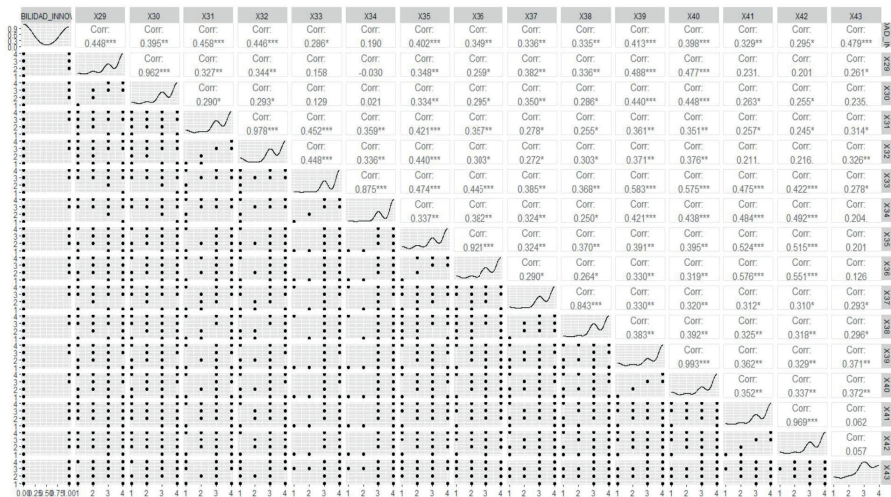
c) Factor de innovación ambiental, variables de innovación entre X12 y X19

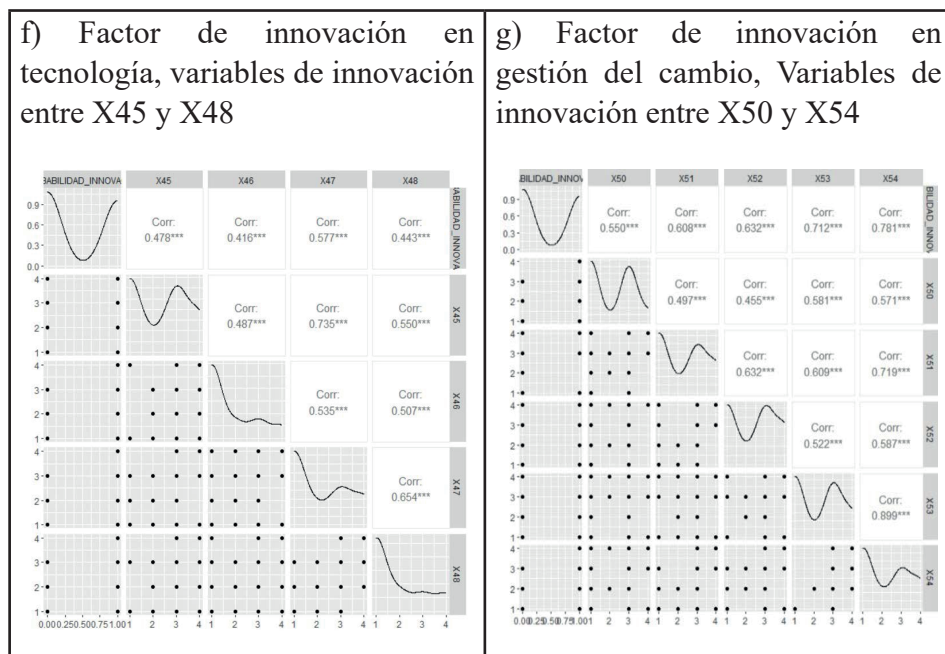


d) Factor de innovación producción, variables de innovación entre X21 y X27



e) Factor de innovación conocimiento, variables de innovación entre X29 y X43





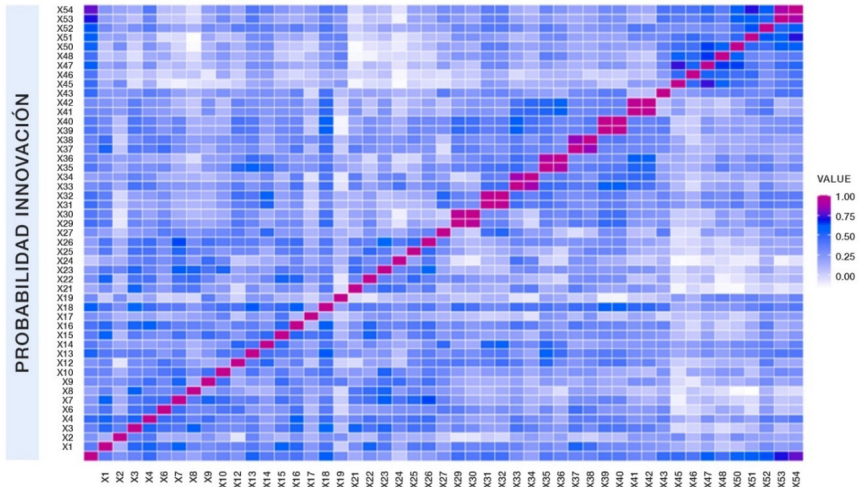
Nota. Elaboración propia con base en el procesamiento de datos mediante el software R-Commander.

El análisis de los resultados identifica variables clave que incrementan de forma significativa la probabilidad de innovación en los procesos productivos de las fincas, fortaleciendo su capacidad para obtener café diferenciado. Además, se identifican correlaciones estadísticamente significativas (***) entre varias variables asociadas a los factores de innovación, lo que plantea la necesidad de emplear herramientas adicionales que permitan seleccionar con mayor precisión las variables con mayor probabilidad de incidir en la innovación, orientadas a mejorar los procesos productivos que posibiliten la generación de café de especialidad.

6.3.2 Resultados del análisis correlacional

El análisis correlacional se realizó mediante mapas de calor (Figura 19). La figura permite identificar qué variables independientes presentan una relación más alta con la variable dependiente probabilidad innovación, con la finalidad de identificar procesos que permita generar un producto diferenciado.

Figura 19. Mapa de calor de correlaciones entre variables de innovación



Nota. Elaboración propia con base en el procesamiento de datos mediante el software R-Commander.

En la tabla 3 se presentan los datos de correlación entre las variables y la probabilidad de innovación para que una finca cafetera pueda mejorar los procesos de producción para producir café especial de tipo sostenible.

Tabla 3. Valores de correlación entre variables de innovación con la probabilidad de innovación

Variabes probabilidad de innovación	Valores de correlación	Variabes probabilidad de innovación	Valores de correlación
PROBABILIDAD_INNOVACION	1	X29	0.44792566
X1	0.37555102	X30	0.39457241
X2	0.26316289	X31	0.45769961
X3	0.33692435	X32	0.446139
X4	0.47731394	X33	0.28631919
X6	0.26992734	X34	0.19028141
X7	0.24230407	X35	0.40245135
X8	0.17477374	X36	0.34892755

VARIABLES PROBABILIDAD DE INNOVACIÓN	VALORES DE CORRELACIÓN	VARIABLES PROBABILIDAD DE INNOVACIÓN	VALORES DE CORRELACIÓN
X9	0.29287039	X37	0.33572222
X10	0.33789592	X38	0.33540808
X12	0.25859463	X39	0.41325396
X13	0.53129045	X40	0.3983849
X14	0.38501001	X41	0.32855062
X15	0.37555102	X42	0.29540944
X16	0.42238031	X43	0.47932973
X17	0.27551925	X45	0.47825784
X18	0.58851698	X46	0.41582195
X19	0.29724095	X47	0.57697456
X21	0.14368424	X48	0.44257688
X22	0.27664217	X50	0.54957537
X23	0.22848131	X51	0.60752826
X24	0.03588917	X52	0.63170346
X25	0.20033972	X53	0.71181891
X26	0.20654211	X54	0.7808601
X27	0.28503131		

Nota. Elaboración propia con base en el procesamiento de datos mediante el software R-Commander.

Los resultados del análisis correlacional revelaron diez variables con coeficientes de correlación de Pearson superiores a 0.47, consideradas como correlaciones fuertes:

1. **X54** (Gestión del cambio): 0.7808601
2. **X53** (Gestión del cambio): 0.71181891
3. **X52** (Gestión del cambio): 0.63170346
4. **X51** (Gestión del cambio): 0.60752826
5. **X18** (Ambiental): 0.58851698

6. **X47** (Tecnología): 0.57697456
7. **X50** (Gestión del cambio): 0.54957537
8. **X13** (Ambiental): 0.53129045
9. **X43** (Conocimiento): 0.47932973
10. **X4** (Económico): 0.47731394

Este análisis confirmó la importancia primaria de los factores de gestión del cambio, ambientales y tecnológicos como determinantes de la capacidad innovadora de las fincas cafeteras.

6.4 Selección del método estadístico para el modelo

6.4.1 Evaluación de alternativas

La naturaleza de los datos y el objetivo predictivo del modelo orientaron la evaluación de diferentes métodos estadísticos. Se consideraron principalmente dos alternativas:

1. Regresión lineal: Inicialmente evaluada por su simplicidad interpretativa, pero descartada porque:

- Asume una variable dependiente continua, mientras que la probabilidad de innovación es binaria (0/1).
- Puede predecir valores fuera del rango 0-1, inadecuados para una probabilidad.
- No maneja adecuadamente la relación no lineal entre variables independientes y la probabilidad de ocurrencia.

2. Regresión logística: Seleccionada como método óptimo debido a:

Esta técnica estudia la relación entre una variable dependiente y múltiples variables independientes sean numéricos o por categorías.

- No hay una relación directa entre las variables dependientes e independientes.
- Proporciona odds ratios fácilmente interpretables.
- Maneja adecuadamente variables independientes ordinales (medidas en escala Likert de 1-4).

6.4.2 Fundamentos de la regresión logística

La regresión logística es un método estadístico que calcula la probabilidad de que ocurra un evento en función de un conjunto de variables. El modelo se basa en la transformación logit, que convierte una probabilidad (restringida entre 0 y 1) en una escala logarítmica que puede variar de menos infinito a más infinito. La ecuación (2) fundamental del modelo se expresa como (Zou et al., 2019):

$$\ln(\pi/(1 - \pi)) = \text{Beta}_0 + \text{Beta}_1 * X_1 + \dots + \text{Beta}_k * X_k \quad (2)$$

Donde:

- π indica la probabilidad de que una finca sea innovadora o no innovadora.
- $\ln(\pi/(1-\pi))$ es el logaritmo natural de las probabilidades a favor del evento.
- beta_0 es el término constante o intercepto.
- $\text{beta}_1, \text{beta}_2, \dots, \text{beta}_k$ son los coeficientes asociados a las variables independientes X_1, X_2, \dots, X_k .

Este método resulta particularmente adecuado para el objetivo de predecir la probabilidad de innovación en fincas cafeteras, permitiendo determinar qué variables tienen mayor influencia en esta probabilidad y cuantificar su impacto.

6.5 Desarrollo del modelo de regresión logística

6.5.1 Análisis de variables mediante ANOVA

El análisis ANOVA se realizó de manera sistemática para cada una de las 54 variables de innovación identificadas, agrupadas según los siete factores de innovación. La tabla 4 presenta un resumen del análisis por factor, para cada

variable se calculó el estadístico Chi-cuadrado y su valor de significancia (p-value), estableciendo como criterio de selección $p < 0.05$ (nivel de confianza del 95 %).

Tabla 4 *Análisis ANOVA con las todas variables de innovación*

Factor de innovación	Variables analizadas	Significativas ($p < 0.05$)	No significativas	Variables destacadas
F1: Económico	5	1	4	X4 ($p=0.100$)
F2: Social	6	0	6	Ninguna
F3: Ambiental	9	1	8	X18 ($p < 0.001$)***
F4: Producción	8	0	8	Ninguna
F5: Conocimiento	16	1	15	X43 ($p=0.019$)**
F6: Tecnología	5	1	4	X47 ($p < 0.001$)***
F7: Gestión cambio	5	3	2	X50, X51***, X52**
TOTAL	54	7	47	—

Nota. Elaboración propia con base en el procesamiento de datos mediante el software R-Commander.

Los resultados revelan que solo 7 (13 %) presentan diferencias estadísticamente significativas ($p < 0.05$) entre fincas innovadoras que producen café especial y normal. Es notable que 47 variables (87 %) no mostraron diferencias significativas, lo que indica que la innovación en fincas cafeteras no se caracteriza por cambios en todas las variables, sino por transformaciones específicas en áreas críticas

Con las 7 variables significativas identificadas, se aplicó un proceso de selección en tres etapas para determinar el conjunto óptimo de predictores. La tabla 5 presenta los resultados de cada análisis:

Análisis 1, se evalúan las 7 variables significativas conjuntamente. Los resultados mostraron que cuatro variables (X18, X43, X47, X51) mantuvieron alta significancia, mientras que X4, X50 y X52 mostraron significancia reducida en presencia de las otras.

Análisis 2, se afina el modelo con las 4 variables seleccionadas. Tres de ellas

(X18, X47, X51) presentaron los valores p más bajos (< 0.001) y estadísticos Chi-cuadrado más altos.

Análisis 3: Se validó el modelo final con las 3 variables seleccionadas, confirmando que X18, X47 y X51 mantienen alta significancia ($p < 0.001$) cuando se analizan conjuntamente.

Tabla 5. *Análisis ANOVA de variables con mayor impacto*

Análisis 1					
Factor de innovación	Variables	Chisq	Df	Pr(>Chisq)	Nivel Significativo
Económico	X4	2.6908	1	0.1009304	
Ambiental	X18	4.4618	1	0.0346601	*
Conocimiento	X43	4.1823	1	0.0408482	*
Tecnología	X47	13.3778	1	0.0002546	***
Gestión del cambio	X50	2.567	1	0.1091171	
Gestión del cambio	X51	5.6509	1	0.017446	*
Gestión del cambio	X52	3.4791	1	0.0621477	.
Análisis 2					
Factor de innovación	Variables	Chisq	Df	Pr(>Chisq)	Nivel Significativo
Ambiental	X18	13.345	1	0.0002591	***
Conocimiento	X43	3.981	1	0.0460159	*
Tecnología	X47	17.634	1	0.000026781	***
Gestión del cambio	X51	19.863	1	0.000008321	***
Análisis 3					
Factor de innovación	Variables	Chisq	Df	Pr(>Chisq)	Nivel Significativo
Ambiental	X18	19.451	1	0.000010322	***
Tecnología	X47	15.16	1	0.000098775	***
Gestión del cambio	X51	19.758	1	0.000008789	***

Nota. Elaboración propia con base en el procesamiento de datos mediante el software R-Commander.

Los resultados del análisis ANOVA confirmaron la significancia estadística de tres variables principales:

- **X18 (Gestión integral de residuos)**: variable del factor ambiental con un valor p de 0.000010322, altamente significativo.
- **X47 (Adopción de tecnología digital)**: variable del factor tecnológico con un valor p de 0.000098775, también altamente significativo.
- **X51 (Flexibilidad ante el cambio)**: variable del factor de gestión del cambio con un valor p de 0.000008789, mostrando la mayor significancia estadística.

Estas tres variables fueron seleccionadas como los predictores claves para el desarrollo del modelo de regresión logística.

6.5.2 Construcción del modelo

Con base en las tres variables identificadas como más significativas mediante el análisis ANOVA y correlacional (X18 - Gestión integral de residuos, X47 - Adopción de tecnología digital, y X51 - Flexibilidad ante el cambio), se procedió a la construcción del modelo de regresión logística mediante el software R-Commander.

La especificación del modelo se basó en los siguientes criterios metodológicos: a) Variable dependiente binaria (Y) codificada como 1 para fincas innovadoras (productoras de café especial) y 0 para fincas con procesos de producción normal, b) Variables independientes ordinales medidas en escala Likert de 1 a 4, representando el nivel de implementación de cada práctica, c) Muestra de 66 observaciones (31 fincas innovadoras y 35 fincas tradicionales), d) Análisis realizado mediante el software R-Commander, herramienta que facilita el análisis estadístico con interfaz gráfica para regresión logística.

El proceso de construcción del modelo siguió una secuencia sistemática: Primero, se evaluó un modelo de regresión logística incluyendo todas las variables de los factores de innovación identificadas mediante el análisis ANOVA (X4, X18, X43, X47, X50, X51, X52). Segundo, se identificó que algunas variables perdían significancia en presencia de otras, por lo que se afinó el modelo con las 4 variables que mantuvieron significancia consistente (X18, X43, X47, X51). Tercero, tras evaluar los coeficientes estimados mediante R-Commander y la capacidad predictiva de cada variable, se determinó que

X43 presentaba un aporte marginal menor, resultando en el modelo final con 3 variables (X18, X47, X51) que demostraron la mayor significancia estadística y capacidad predictiva. Este proceso permitió identificar un modelo que equilibra simplicidad con capacidad explicativa. La tabla 6 presenta los coeficientes estimados del modelo final.

Tabla 6. *Coefficientes del modelo de regresión logística*

Descripción	Estimado	Std. Error	z value	Pr(> z)	Nivel significativo
Intercepto	-26.6951	7.939	-3.363	0.000772	***
X18	4.861	1.5995	3.039	0.002373	**
X47	2.1631	0.7833	2.762	0.005752	**
X51	2.0838	0.7098	2.936	0.003328	**

Nota. Elaboración propia con base en el procesamiento de datos mediante el software R-Commander.

La validación del modelo se fundamenta en un análisis que abarca diferentes dimensiones. El análisis ANOVA permitió identificar las variables X18 (gestión integral de residuos), X47 (adopción de tecnología digital) y X51 (flexibilidad ante el cambio) como las más significativas, proporcionando una base sólida para el modelo. La regresión logística confirma la relevancia de estas variables, con coeficientes estadísticamente significativos ($p < 0.05$) para todas ellas. La verificación de los supuestos paramétricos, incluyendo normalidad, homogeneidad, linealidad e independencia a través de los gráficos de residuos, demostró que el modelo cumple con los requisitos estadísticos fundamentales.

Finalmente, la evaluación de la capacidad predictiva mediante la curva ROC arrojó un área bajo la curva (AUC) de 0.97 (IC 95 %: 0,945 – 1.0), indicando una excelente capacidad discriminativa del modelo con una precisión del 97 % en la clasificación de fincas cafeteras. Esta convergencia de evidencias estadísticas respalda la robustez y validez del modelo.

La ecuación (3) del modelo de regresión logística resultante es:

$$Y \text{ (Probabilidad de innovación)}$$

$$= -26.6951 + 4.861 \times X18 + 2.1631 \times X47 + 2.0838 \times X51 \quad (3)$$

Donde:

- Y representa la probabilidad de que una finca sea clasificada como innovadora.
- X18 corresponde al nivel de implementación de gestión integral de residuos.
- X47 representa el grado de adopción de tecnología digital.
- X51 indica el nivel de flexibilidad ante el cambio.

Este modelo permite calcular la probabilidad de que una finca sea innovadora (y por tanto capaz de producir café especial sostenible) basándose en la evaluación de estas tres variables clave.

6.5.3 Interpretación de los coeficientes

Los coeficientes del modelo proporcionan información valiosa sobre la influencia relativa de cada variable:

- **X18 (Gestión integral de residuos):** Con un coeficiente de 4.861, esta variable muestra el mayor impacto en la probabilidad de innovación. Por cada mejora en sus prácticas de gestión en residuos, las probabilidades de que una finca sea innovadora aumentan en 4.861 unidades.
- **X47 (Adopción de tecnología digital):** Con un coeficiente de 2.1631, esta variable tiene un impacto considerable. Cada unidad de incremento en la adopción digital aumenta el logaritmo de las probabilidades en 2.1631 unidades.
- **X51 (Flexibilidad ante el cambio):** Con un coeficiente de 2.0838, esta variable muestra un impacto similar a la adopción digital. Cada unidad de incremento en la flexibilidad aumenta el logaritmo de las probabilidades en 2.0838 unidades.

El intercepto negativo (-26.6951) indica que, en ausencia de las tres variables críticas (o con valores muy bajos), la probabilidad de que una finca sea innovadora es prácticamente nula.

6.6 Descripción detallada de las variables clave

6.6.1 Gestión integral de residuos (X18)

Esta variable, perteneciente al factor ambiental, evalúa la capacidad de la finca para identificar y gestionar de manera segura los diferentes tipos de residuos generados durante el proceso productivo. Una implementación óptima de esta variable incluye:

- **Identificación y categorización:** Clasificación sistemática de los diferentes tipos de residuos (orgánicos, inorgánicos, peligrosos) generados en cada etapa del proceso.
- **Prácticas de segregación:** Implementación de sistemas para separar residuos desde su origen, facilitando su manejo posterior.
- **Maximización de reutilización y reciclaje:** Desarrollo de prácticas que permitan aprovechar residuos como recurso, particularmente la pulpa de café como abono y los líquidos para la producción de gases.
- **Disposición adecuada:** Establecimiento de procedimientos seguros para la eliminación de residuos no recuperables, minimizando impactos ambientales.
- **Áreas específicas:** Creación de zonas designadas para almacenamiento y procesamiento de residuos, evitando contaminación cruzada.

La gestión integral de residuos no solo minimiza impactos negativos en salud y medio ambiente, sino que también genera valor agregado al transformar lo que tradicionalmente se consideraba desecho en recurso productivo.

6.6.2 Adopción de tecnología digital (X47)

Esta variable, componente del factor tecnológico, evalúa el grado de digitalización en la gestión de información relacionada con los procesos productivos, calidad, capacitaciones y aspectos económicos. Una implementación avanzada de esta variable comprende:

- **Digitalización documental:** Transformación de registros físicos a formatos digitales accesibles y seguros.

- **Sistemas de trazabilidad:** Implementación de herramientas digitales para seguimiento del producto en cada etapa del proceso de producción de las fincas cafeteras.
- **Gestión de datos productivos:** Registro sistemático de variables productivas (rendimientos, calidades, insumos) en formatos que permitan análisis posterior.
- **Comunicación digital:** Uso de plataformas digitales para conectar con compradores, certificadoras y comunidades de práctica.

Respuesta a cambios de mercado: Utilización de herramientas digitales para monitorear tendencias y adaptar estrategias comerciales.

- La adopción digital es un camino ordenado para dar uso y manejar la información con la finalidad de decidir mejor, la cual puede estar implementado en cada proceso de producción.

6.6.3 Flexibilidad ante el cambio (X51)

Esta variable, parte del factor de gestión del cambio, evalúa la disposición y capacidad de la finca para adaptarse a modificaciones en el ambiente productivo y comercial. Una implementación efectiva incluye:

- **Adaptación a nuevos roles:** Capacidad para redistribuir responsabilidades según necesidades emergentes.
- **Modificación de procesos:** Disposición para ajustar procedimientos en respuesta a nueva información o circunstancias.
- **Adopción de nuevos métodos:** Apertura para incorporar técnicas y prácticas innovadoras que mejoren resultados.
- **Comunicación abierta:** Desarrollo de canales efectivos para compartir información y recibir retroalimentación.
- **Gestión de sugerencias:** Implementación de mecanismos para que trabajadores y colaboradores puedan proponer mejoras.

La flexibilidad ante el cambio representa un factor crítico en un entorno caracterizado por fluctuaciones de mercado, variabilidad climática y evolución constante de estándares y certificaciones.

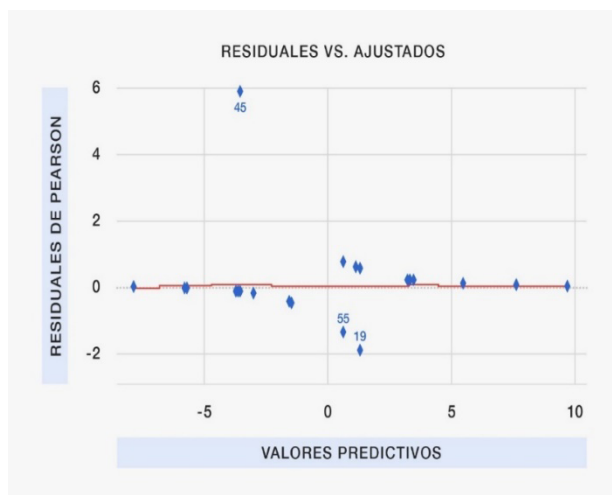
6.7 Validación del modelo

6.7.1 Verificación de supuestos estadísticos

La validación del modelo requirió verificar el cumplimiento de los supuestos fundamentales de la regresión logística:

- **Linealidad:** verifica la relación entre las variables independientes y la probabilidad de innovación mediante análisis de residuos. Como se puede observar en la figura 20, los residuos están esparcidos aleatoria sin que presente patrones evidentes, confirmando el cumplimiento del supuesto de linealidad.

Figura 20. Residuos versus valores ajustados



Nota. Elaboración propia con base en el procesamiento de datos mediante el software R-Commander.

- **Independencia de observaciones:** El diseño muestral mediante selección aleatoria simple aseguró la independencia entre observaciones.

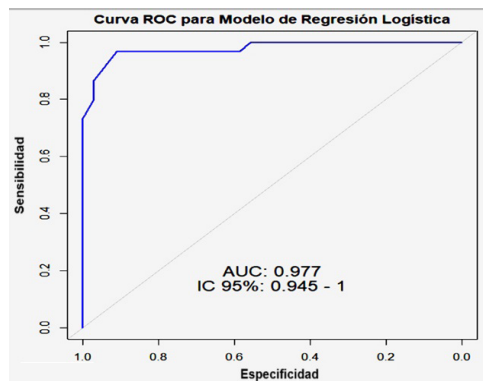
- **Ausencia de multicolinealidad:** Se verificó que las variables predictoras no presentaran correlaciones excesivamente altas entre sí (todas por debajo de 0.7), evitando problemas de multicolinealidad.
- **Ausencia de valores atípicos influyentes:** El análisis de distancias de Cook y gráficos de apalancamiento permitió identificar la observación 45 como potencialmente atípica. Sin embargo, su distancia de Cook permanecía por debajo del umbral crítico de 1, y análisis de sensibilidad excluyendo esta observación confirmaron que no afectaba significativamente los coeficientes del modelo.
- **Tamaño muestral adecuado:** La regla general de al menos 10 eventos (casos positivos) por variable predictoras se cumplió adecuadamente, con 31 fincas clasificadas como innovadoras y 3 predictores en el modelo final.

El cumplimiento satisfactorio de estos supuestos confirma la validez estadística del modelo desarrollado.

6.7.2 Evaluación de la capacidad predictiva

Para evaluar el poder predictivo del modelo de innovación, se desarrolló mediante la curva ROC (Receiver Operating Characteristic) y calculando el área bajo la curva (AUC). La figura 21 muestra la curva ROC del modelo.

Figura 21. Curva ROC para evaluar la capacidad predictiva del modelo



Nota. Elaboración propia con base en el procesamiento de datos mediante el software R-Commander.

El resultado obtenido del área bajo la curva ROC es de 0.97 (IC 95 %), lo que indica una excelente capacidad discriminativa del modelo. Este valor significa que el modelo tiene un 97 % de probabilidad de clasificar correctamente un par de fincas (una innovadora y otra no innovadora) basándose en las tres variables predictoras identificadas.

6.7.3 Análisis de sensibilidad

Se realizó un análisis de sensibilidad para comprar la capacidad predictiva del modelo de innovación versus los datos obtenidos en campo las fincas que tiene procesos de innovación para generar café diferenciado. La tabla 7 muestra los resultados acordes a las predicciones del modelo de innovación para una muestra de 20 fincas.

Tabla 7. Análisis de sensibilidad del modelo

Resultados de encuestas					Resultados ecuación del modelo de innovación					Acierto
Finca	X17	X48	X51	Tipo de Café	X18	X47	X51	Probabilidad	Tipo de Café	Si/No)
1	3	1	3	Normal	3	1	3	0.02418359	Normal	Si
2	3	1	3	Normal	3	1	3	0.02418359	Normal	Si
3	3	2	3	Normal	3	2	3	0.17733624	Normal	Si
4	3	2	3	Especial	3	2	3	0.17733624	Normal	No
5	4	1	4	Especial	4	1	4	0.96257237	Especial	Si
6	4	1	1	Normal	4	1	1	0.04723649	Normal	Si
7	4	3	4	Especial	4	3	4	0.99948632	Especial	Si
8	1	1	1	Normal	1	1	1	2.3013E-08	Normal	Si
9	4	1	1	Normal	4	1	1	0.04723649	Normal	Si
10	4	4	4	Especial	4	4	4	0.99994092	Especial	Si
11	4	1	1	Normal	4	1	1	0.04723649	Normal	Si
12	4	1	1	Normal	4	1	1	0.04723649	Normal	Si
13	4	1	3	Especial	4	1	3	0.76194996	Especial	Si
14	4	1	4	Especial	4	1	4	0.96257237	Especial	Si

Resultados de encuestas					Resultados ecuación del modelo de innovación					Acierto
Finca	X17	X48	X51	Tipo de Café	X18	X47	X51	Probabilidad	Tipo de Café	Si/No)
15	4	1	1	Normal	4	1	1	0.04723649	Normal	Si
16	4	2	4	Especial	4	2	4	0.9955496	Especial	Si
17	4	1	1	Normal	4	1	1	0.04723649	Normal	Si
18	4	1	1	Normal	4	1	1	0.04723649	Normal	Si
19	4	3	4	Especial	4	3	4	0.99948632	Especial	Si
20	4	2	3	Especial	4	2	3	0.96532681	Especial	Si

Nota. Elaboración propia.

El análisis de sensibilidad reveló que el modelo predice correctamente el tipo de finca (innovadora/no innovadora) en 19 de los 20 casos analizados, lo que representa una precisión del 95 %. El único caso de diferencia correspondía a una finca clasificada como productora de café especial, pero con valores en las variables predictoras más alineados con fincas no innovadoras, sugiriendo posibles inconsistencias en la recolección de datos original.

Esta alta tasa de acierto confirma la robustez del modelo y su aplicabilidad práctica como herramienta predictiva en el contexto de fincas cafeteras.

6.8 Aplicaciones prácticas del modelo

6.8.1 Herramienta diagnóstica

El modelo desarrollado puede utilizarse como herramienta diagnóstica para evaluar el potencial innovador de una finca cafetera, siguiendo estos pasos:

- **Evaluación inicial:** calificar en escala de 1 a 4 el nivel de implementación de cada una de las tres variables clave:
 - ★ Gestión integral de residuos (X18).
 - ★ Adopción digital (X47).
 - ★ Gestión al cambio (X51).

- **Cálculo de probabilidad:** aplicar la ecuación del modelo para obtener la probabilidad de que la finca sea innovadora: Y (Probabilidad de innovación) = $-26.6951 + 4.861 * X_{18} + 2.1631 * X_{47} + 2.0838 * X_{51}$
- **Interpretación:** una probabilidad mayor a 0.5 sugiere que la finca tiene potencial para ser innovadora y producir café especial; valores superiores a 0.7 indican alto potencial innovador.
- **Identificación de brechas:** comparar los valores de las tres variables clave con los valores óptimos (4 en la escala) para identificar áreas de mejora prioritarias.

Esta aplicación diagnóstica permite obtener una evaluación objetiva del potencial innovador de una finca y orientar los esfuerzos de mejora hacia las variables más influyentes.

6.8.2 Guía para la intervención

El modelo también funciona como guía para diseñar intervenciones efectivas que potencien la innovación en fincas cafeteras. Los coeficientes del modelo revelan el impacto relativo de cada variable, orientando la priorización de acciones:

- **Gestión de residuos** (coeficiente 4.861): como la variable de mayor impacto, las intervenciones deberían comenzar por implementar o mejorar los sistemas de identificación, clasificación y aprovechamiento de residuos.
- **Adopción digital** (coeficiente 2.1631): en segundo orden de prioridad, las intervenciones deberían facilitar la implementación de herramientas digitales básicas para gestión de información.
- **Flexibilidad al cambio** (coeficiente 2.0838): como tercer elemento, las intervenciones deberían fomentar la apertura a nuevas prácticas y métodos, principalmente a través de procesos sensibilización y capacitación.

Esta priorización permite optimizar recursos limitados, dirigiéndolos a las variables con mayor potencial de impacto en la capacidad innovadora.

6.8.3 Instrumento de monitoreo

El modelo proporciona un marco sistemático para monitorear la evolución de fincas en procesos de transformación hacia la producción de café especial:

- **Evaluación periódica:** aplicar la herramienta diagnóstica en intervalos definidos (semestrales o anuales) para cuantificar avances.
- **Seguimiento de transición:** monitorear cómo evoluciona la probabilidad calculada a medida que se implementan mejoras en las variables clave.
- **Verificación de impactos:** correlacionar los cambios en la probabilidad de innovación con mejoras efectivas en calidad del café y precios obtenidos.

Este enfoque permite ajustar estrategias de intervención según su efectividad demostrada y ofrece evidencia objetiva de progreso para caficultores y entidades de apoyo.

6.9. Limitaciones del modelo

6.9.1 Alcance geográfico y contextual

El modelo fue desarrollado con datos de fincas cafeteras de las veredas Guacacallos y Acacos del municipio de Pitalito, Huila. Aunque estas áreas son representativas de la producción cafetera colombiana, existen algunas limitaciones en su generalización:

- **Especificidad regional:** las condiciones agroecológicas, socioeconómicas y culturales específicas de Pitalito pueden no ser directamente transferibles a otras regiones cafeteras con diferentes características.
- **Temporalidad:** los datos fueron recolectados en octubre de 2023, reflejando la situación específica de ese momento. Factores estacionales o tendencias a largo plazo podrían no estar completamente capturados.
- **Enfoque en café arábica:** el estudio se centró exclusivamente en productores de café arábica, limitando su aplicabilidad a otros tipos de café.

La aplicación del modelo en otros contextos debe considerar estas limitaciones y de ser necesario, ajustar algunos parámetros según las condiciones locales específicas.

6.9.2 Limitaciones metodológicas

Algunas limitaciones metodológicas deben considerarse al aplicar el modelo:

Simplificación de la realidad: como todo modelo, representa una simplificación de la realidad compleja de las fincas cafeteras. Aunque las tres variables seleccionadas tienen alto poder predictivo, existen otros factores que pueden tener alguna injerencia en el modelo que afecte su capacidad predictiva.

Naturaleza estática: el modelo captura relaciones en un momento específico, mientras que la innovación es un proceso dinámico que evoluciona con el tiempo.

Subjetividad en medición: la evaluación de variables como flexibilidad al cambio implica cierto grado de subjetividad, lo que puede afectar la consistencia entre diferentes evaluadores.

Tamaño muestral: aunque suficiente para los análisis realizados, un tamaño muestral mayor podría haber permitido identificar relaciones más sutiles o variables adicionales relevantes.

Estas limitaciones no invalidan el modelo, pero deben considerarse en su interpretación y aplicación.

6.9.3 Consideraciones para aplicación práctica

Al implementar el modelo en contextos reales, deben considerarse algunos factores adicionales:

- **Retroalimentación iterativa:** el modelo debe entenderse como un punto de partida que puede refinarse con retroalimentación de su aplicación en campo.
- **Complementariedad:** los resultados del modelo deben complementarse con conocimiento contextual y experiencia práctica de técnicos y caficultores.

- **Adaptabilidad:** la importancia relativa de las variables puede variar según características específicas de la finca y condiciones locales.
- **Enfoque integral:** aunque el modelo identifica tres variables críticas, la innovación efectiva requiere un enfoque sistémico que considere múltiples dimensiones de la finca cafetera.

El reconocimiento de estas consideraciones favorece una aplicación más efectiva y contextualmente apropiada del modelo.

Conclusiones

El análisis de los datos identifica que la variable X18 de gestión integral de residuos tiene el mayor impacto (coeficiente 4.861), seguida por la variable X47 adopción digital (2.1631) y por último la variable X51 de flexibilidad al cambio (2.0838). Estos resultados permiten que los caficultores se enfoquen en los procesos importantes, para alinearse en las variables que les permitirá generar un producto diferencial.

La metodología desarrollada en la investigación por medio de análisis de clúster, mapas de calor, análisis ANOVA y validación paramétrica, proporciona una base teórica para el desarrollo de modelos de innovación similares en otros contextos agrícolas o cafeteras. Estas técnicas de análisis multivariado y de validación cruzada aseguran que el documento tenga una solidez científica con resultados confiables.

El modelo de innovación va más allá del ámbito académico al proporcionar una herramienta directamente aplicable por caficultores, técnicos, cooperativas y entidades gubernamentales. Su aplicación como instrumento de diagnóstico, guía de intervención y sistema de monitoreo de procesos productivos representa un aporte práctico alto valor para la promoción de la innovación en el sector cafetero.

Las limitaciones identificadas y las consideraciones para su aplicación práctica establecen lineamientos claros para futuras investigaciones orientadas a mejorar, adaptar o expandir el modelo a otros contextos agrícolas sean cafeteros u orientados a otros productos.

