

SEMILLA: Inteligencia territorial para anticipar riesgos agroclimáticos y alimentarios

1. Resumen

SEMILLA es una solución de inteligencia territorial diseñada para anticipar, priorizar y explicar los riesgos agroclimáticos y alimentarios en los 1.122 municipios de Colombia. La herramienta integra datos abiertos de diversas fuentes, imágenes satelitales (Sentinel-2) e inteligencia artificial para transformar información territorial compleja en un **KPI municipal de riesgo**.

El proyecto aborda la problemática de la información dispersa y fragmentada, que impide a las entidades públicas responder eficazmente a preguntas críticas sobre seguridad alimentaria y resiliencia rural. SEMILLA supera esta limitación mediante un análisis multidimensional que combina la exposición climática, la vulnerabilidad social, la resiliencia ecológica y la capacidad de respuesta de cada territorio.

El resultado principal es una herramienta de consulta pública que permite: (1) priorizar municipios según su nivel de riesgo, (2) identificar perfiles territoriales mediante clustering y (3) orientar intervenciones diferenciadas en seguridad alimentaria, asistencia técnica, conectividad rural y adaptación climática. La solución se enmarca en el desafío de implementar modelos de IA para predecir rendimientos agrícolas y riesgos climáticos, demostrando un impacto potencial en la productividad y resiliencia de las comunidades rurales.

2. Descripción del Problema

La variabilidad y el cambio climático representan una amenaza creciente para la producción agrícola, la disponibilidad de alimentos y los medios de vida en las zonas rurales de Colombia. Sin embargo, las entidades del sector público y los tomadores de decisiones a menudo carecen de una visión integrada que les permita:

1. Identificar municipios que combinan alta exposición climática con alta vulnerabilidad alimentaria.
2. Evaluar cómo la reducción de la cobertura vegetal afecta la productividad y la seguridad alimentaria local.
3. Reconocer las brechas de conectividad vial que dificultan la asistencia oportuna en emergencias.
4. Agrupar territorios con características similares para diseñar intervenciones diferenciadas y eficientes.
5. Priorizar la asignación de recursos públicos para prevenir el deterioro nutricional y fortalecer la resiliencia.

SEMILLA resuelve este problema consolidando información de múltiples fuentes en una plataforma única que convierte datos complejos en conocimientos claros, explicables y utilizables para la gestión pública.

3. Objetivos

Objetivo General:

Desarrollar una solución basada en datos abiertos, inteligencia artificial, Machine Learning, que permita anticipar, priorizar y explicar riesgos territoriales para la toma de decisiones públicas, y de prevención frente a agroclimáticos y seguridad alimentarios a nivel municipal en Colombia.

Objetivos Específicos:

1. Integrar fuentes abiertas de información agrícola, climática, territorial, vial, poblacional, satelital y social a escala municipal.
2. Construir variables comparables para caracterizar la exposición climática, resiliencia ecológica, vulnerabilidad alimentaria, conectividad y ruralidad.
3. Diseñar un KPI municipal de riesgo territorial que permita priorizar municipios.
4. Identificar grupos de municipios con características similares mediante clustering.
5. Incorporar análisis predictivo y modelos espacio-temporales para anticipar tendencias de riesgo.
6. Presentar los resultados en una página web de consulta clara y orientada a usuarios públicos.

4. Fuentes de Datos

El proyecto integra información proveniente de fuentes públicas y abiertas, garantizando la replicabilidad y actualización del análisis.

Fuente	Tipo de Información	Uso en SEMILLA
UPRA	Producción agrícola, cultivos, uso del suelo	Caracterización productiva y diversidad agrícola
DANE	Población, ruralidad, inseguridad alimentaria	Medición de exposición social y vulnerabilidad alimentaria
Instituto Nacional de Salud	Salud, nutrición y variables de riesgo social	Aproximación al impacto social y nutricional
IGAC	Capas geográficas y territoriales	Integración espacial y caracterización municipal
OpenStreetMap	Red vial y accesibilidad terrestre	Construcción de indicadores de suficiencia vial
Open-Meteo	Temperatura, precipitación y variables climáticas	Análisis de exposición climática
Sentinel-2 / GEE	Índices satelitales (NDVI, NDMI, etc.)	Medición de cobertura vegetal y cambios territoriales

5. Integración de Fuentes y Construcción de Variables Clave

La construcción de un modelo de riesgo territorial robusto y explicable requiere la integración armoniosa de múltiples dimensiones. El proyecto SEMILLA consolida información de diversas fuentes de datos abiertos para crear un conjunto de variables comparables a escala municipal. Este proceso de integración no solo se enfoca en la disponibilidad de datos, sino en su relevancia conceptual para capturar los factores que determinan la transición de un choque climático a una crisis alimentaria y nutricional.

5.1. Proceso de Integración y Homologación

Toda la información fue integrada utilizando los códigos DIVIPOLA del DANE como estándar de homologación territorial. El proceso consistió en:

1. **Recopilación:** Descarga de datos desde las plataformas oficiales de cada fuente.
2. **Limpieza:** Depuración de valores atípicos, manejo de datos faltantes mediante imputación y estandarización de formatos.
3. **Homologación:** Asignación de cada registro al código municipal correspondiente, asegurando una escala de análisis consistente para los 1.122 municipios.
4. **Agregación y Cálculo:** Para variables como las climáticas o satelitales, se calcularon estadísticos (media mensual) para resumir la información a nivel municipal.

5.2. Variables utilizadas y su relación con la inseguridad alimentaria

Cada variable fue seleccionada por su capacidad para representar un eslabón en la cadena causal con la inseguridad alimentaria. A continuación, se detalla la fuente, la construcción y la justificación conceptual de cada una.

- **Métricas de Cobertura Vegetal**

- **Fuente:** Sentinel-2, procesado en Google Earth Engine.
- **Construcción:** Se calcularon diversos índices espectrales (NDVI, EVI, NDMI, NDWI, MSI, BSI, NDBI) para caracterizar el estado de la vegetación, la humedad del suelo y las condiciones de la superficie. Para cada municipio, se calculó la media mensual de estos índices, capturando el vigor, la densidad y la salud de la cobertura vegetal.
- **Justificación:** Las métricas de cobertura vegetal constituyen indicadores fundamentales de la resiliencia ecológica del territorio. Una baja cobertura vegetal se asocia con una mayor susceptibilidad a la erosión, pérdida de capacidad productiva del suelo, inundaciones y sequías. Estas condiciones afectan directamente la producción agropecuaria local, reduciendo la disponibilidad de alimentos y aumentando el riesgo de crisis alimentarias tras eventos climáticos extremos, lo que impacta directamente en el estado nutricional de la población. El análisis econométrico permitió seleccionar el NDVI como el índice más explicativo de la relación con la desnutrición, validando su uso como variable central en el modelo.

- **Índice de Suficiencia Vial**

- **Fuente:** OpenStreetMap.
- **Construcción:** Se calculó el Índice de Engel para cada municipio, una medida que supera el sesgo de la densidad vial al considerar tanto la superficie como la población del territorio. La fórmula utilizada fue:

$$I_e = KmV / \sqrt{(S \times P)}$$

Donde:

I_e = Índice de Engel.

KmV = Longitud en kilómetros de la red vial del municipio (carreteras y vías).

S = Superficie municipal en kilómetros cuadrados.

P = Número de habitantes del municipio.

El Índice de Engel representa la capacidad de la red vial para garantizar los servicios de transporte, considerando la presión demográfica y territorial. Valores más bajos indican una menor capacidad de la infraestructura vial (zonas con redes relativamente saturadas), mientras que valores más altos señalan una mayor capacidad para garantizar la circulación y el intercambio de bienes y personas. Los resultados se agruparon en cinco rangos para facilitar su interpretación.

- **Justificación:** La suficiencia vial es un factor crítico para la articulación territorial y el acceso a bienes y servicios. El trazo de las vías define distintos niveles de articulación territorial a escala local, municipal y regional, determinando la organización y funcionalidad de los principales corredores para el tráfico de pasajeros y mercancías. En el contexto de la seguridad alimentaria y nutricional, la red vial es fundamental por varias razones:

Acceso a mercados y alimentos: Una baja suficiencia vial limita la capacidad de los productores rurales para comercializar sus cosechas y de los hogares para acceder a alimentos variados y nutritivos, especialmente en momentos de escasez local.

Acceso a servicios de salud y protección social: La conectividad vial facilita el acceso a centros de salud, programas de complementación alimentaria y asistencia social, lo cual es crucial para la prevención y atención de la desnutrición.

Respuesta a emergencias: En situaciones de crisis climática o sanitaria, los territorios con baja suficiencia vial enfrentan mayores barreras para recibir ayuda humanitaria, atención nutricional y abastecimiento de alimentos, incrementando el riesgo de deterioro nutricional en la población vulnerable.

Estructuración territorial: La infraestructura vial no solo conecta lugares, sino que estructura las relaciones socioeconómicas a largo plazo. Su incidencia territorial es determinante en la dinámica de desarrollo rural y en la capacidad de adaptación de las comunidades.

- **Estimación de Inseguridad Alimentaria (Moderada o Grave)**

- **Fuente:** DANE (Encuesta FIES).
- **Construcción:** Se utilizó el porcentaje de hogares en el municipio con inseguridad alimentaria moderada o grave, según la Escala de Experiencia de Inseguridad Alimentaria (FIES).
- **Justificación:** Es uno de los predictores más directos del riesgo nutricional. Los hogares con inseguridad alimentaria tienen una capacidad limitada para afrontar choques climáticos y económicos, debido a restricciones en el acceso físico y económico a alimentos suficientes, variados y nutritivos. Un alto nivel de inseguridad alimentaria es un indicador de vulnerabilidad previa que se magnifica ante un evento adverso.

- **Porcentaje de Población Rural**

- **Fuente:** DANE (Censo y proyecciones).
- **Construcción:** Porcentaje de la población municipal que reside en zonas rurales.
- **Justificación:** Las poblaciones rurales dependen en mayor medida de actividades agropecuarias, las cuales son altamente sensibles al clima. Además, suelen enfrentar mayores brechas en el acceso a infraestructura, servicios de salud y programas de protección social. Una alta proporción de población rural puede amplificar los impactos de eventos climáticos sobre la seguridad alimentaria y nutricional local.

- **Diversidad Productiva Agrícola**

- **Fuente:** UPRA (Evaluaciones Agropecuarias Municipales - EVA).
- **Construcción:** Se calculó el número de tipos de cultivos (grupos alimentarios) reportados para el municipio como proxy de diversidad productiva.
- **Justificación:** Esta variable evalúa la capacidad del territorio para garantizar la disponibilidad de alimentos estratégicos. Una menor diversidad productiva, o una concentración en pocos grupos alimentarios, incrementa la vulnerabilidad ante pérdidas por fenómenos climáticos. Vincular este análisis con los grupos alimentarios priorizados por el Plan Nacional de Seguridad Alimentaria y Nutrición (PNSAN) y las Guías Alimentarias Basadas en Alimentos (GABA) permite identificar territorios con fragilidad en su sistema agroalimentario.

- **Variables Climáticas**

- **Fuente:** Open-Meteo (Reanálisis ERA5).

- **Construcción:** Se calcularon estadísticos agregados (medias anuales, extremos) para variables como temperatura y precipitación a escala municipal.
 - **Justificación:** Las variables climáticas son críticas por su alta influencia sobre la productividad agrícola, la disponibilidad hídrica y la ocurrencia de eventos extremos. Incrementos sostenidos de temperatura y variaciones en los patrones de precipitación pueden reducir rendimientos, afectar cultivos estratégicos para la alimentación y favorecer condiciones de estrés hídrico, impactando negativamente los medios de vida rurales y, por ende, la nutrición. El análisis econométrico permitió seleccionar la temperatura como la variable climática con mayor poder explicativo sobre los casos de desnutrición.
- **Índice de Riesgo de la Calidad del Agua para Consumo Humano (IRCA)**
 - **Fuente:** Reportes de entidades de salud y ambiente.
 - **Construcción:** Se utilizó el índice reportado para cada municipio, que mide el nivel de riesgo asociado a la calidad del agua para consumo humano.
 - **Justificación:** El acceso a agua de calidad es fundamental para la salud y la nutrición. Un IRCA alto indica un mayor riesgo de enfermedades de transmisión hídrica, las cuales son un factor determinante en la desnutrición aguda, especialmente en la primera infancia, al interferir con la absorción de nutrientes y aumentar el gasto energético.

5.3. Síntesis Conceptual: El Vínculo con la inseguridad alimentaria

La interacción de estas variables permite identificar territorios donde coinciden condiciones de alta exposición climática, baja capacidad de adaptación, limitaciones de acceso a alimentos, fragilidad en la producción agroalimentaria y deficiencias de conectividad. Cuando estos factores convergen, aumenta la probabilidad de que un evento climático desencadene una cascada de efectos: reducción en el consumo de alimentos, pérdida de ingresos, deterioro sanitario y aumento de enfermedades infecciosas. Estas circunstancias constituyen factores de riesgo reconocidos que pueden derivar hasta en casos de desnutrición aguda, especialmente en la primera infancia.

Al focalizar los territorios que presentan la combinación más crítica de estos factores, el índice SEMILLA se convierte en un instrumento de gestión prospectiva que contribuye a la garantía progresiva del Derecho Humano a la Alimentación Adecuada y al fortalecimiento de la resiliencia de los sistemas agroalimentarios territoriales frente a los desafíos del cambio climático.

6. Análisis de Tendencias: Relación Clima-Vegetación y Riesgo Nutricional

Para comprender la dinámica territorial del riesgo, no basta con describir la situación actual; es fundamental analizar las relaciones causales y las tendencias que vinculan las variables climáticas y de cobertura vegetal con la inseguridad alimentaria. En esta etapa, se llevó a cabo un riguroso análisis econométrico utilizando modelos de datos de panel espaciales. El objetivo fue identificar qué variables explican mejor los casos de desnutrición, aproximando así la dimensión nutricional de la seguridad alimentaria. Esto teniendo en cuenta que dada la información disponible para el caso de desnutrición se cuenta con cifras a nivel mensual.

6.1. Selección de Variables y Enfoque Analítico

Se construyó un panel de datos a escala municipal que integra información temporal de diversas fuentes. Las variables dependientes consideradas fueron los casos de desnutrición aguda (reportados por el Ministerio de Salud) como una medida proxy directa de la inseguridad alimentaria y sus consecuencias más severas. Las variables explicativas evaluadas incluyeron un conjunto de indicadores climáticos y satelitales:

- **Climáticas:** Temperatura y precipitación (fuente: Open-Meteo).
- **Satelitales:** Índices derivados de Sentinel-2, procesados en Google Earth Engine, como NDVI, EVI, NDMI, NDWI, MSI, BSI y NDBI.

El enfoque buscó ir más allá de las simples correlaciones para identificar relaciones causales y efectos espaciales que modelaran cómo los cambios en el entorno biofísico influyen en la nutrición de la población.

6.2. Modelos Econométricos Espaciales de Panel

Para abordar la naturaleza longitudinal y espacial de los datos, se implementaron varios modelos de datos de panel. La elección de estos modelos se basó en su capacidad para controlar por la heterogeneidad no observada de los municipios y, crucialmente, para modelar la dependencia espacial. Esta dependencia espacial reconoce que el riesgo en un municipio no es un evento aislado, sino que puede estar influenciado por las condiciones de sus vecinos, fenómeno comúnmente conocido como *contagio espacial*.

Los modelos explorados fueron:

- **Modelo de Poisson con Efectos Fijos:** Dado que la variable dependiente (casos de desnutrición) es un conteo, se partió de un modelo de Poisson. Los efectos fijos permiten controlar por características invariantes en el tiempo de cada municipio, como su geografía o estructura productiva histórica.
- **Modelo de Panel Espacial (splm - Spatial Panel Data Models):** Se utilizó el paquete splm en R, especializado en la estimación de modelos de datos de panel espacial. Esta familia de modelos (SAR, SEM, SDM) permite modelar explícitamente la interacción espacial, ya sea a través de un rezago espacial de la variable dependiente (indicando que la desnutrición en un municipio se propaga a sus vecinos) o a través de un rezago espacial en el término de error (captando factores de riesgo no observados que afectan a grupos de municipios vecinos).
- **Modelo S2SLS (Spatial Two-Stage Least Squares):** Como un enfoque complementario y robusto, se implementó el modelo S2SLS. Este modelo es particularmente útil cuando se sospecha que existe endogeneidad en las variables explicativas, es decir, cuando la relación entre las variables independientes y la dependiente podría estar influenciada por factores no observados que afectan a ambas. El S2SLS aborda este problema utilizando un procedimiento de mínimos cuadrados en dos etapas que incorpora los rezagos espaciales de las variables explicativas como instrumentos. Esto permite obtener estimaciones consistentes del efecto de la temperatura y el NDVI

sobre la desnutrición, incluso en presencia de dependencia espacial y variables endógenas, capturando la inercia y el potencial contagio de los choques climáticos y territoriales entre municipios vecinos. Este modelo fue seleccionado para el análisis final por su capacidad para ofrecer estimaciones robustas en contextos donde las relaciones espaciales son complejas.

6.3. Selección de Variables Finales

Tras la estimación de los modelos, se evaluó la significancia estadística y la consistencia teórica de cada variable explicativa. El análisis permitió destilar el conjunto de variables que presentaban una relación más robusta con los casos de desnutrición:

1. **Temperatura:** Seleccionada como la variable climática más influyente. Los modelos mostraron consistentemente que incrementos en la temperatura media anual están asociados con un aumento significativo en los casos de desnutrición. Este resultado evidencia la sensibilidad de la seguridad alimentaria y la salud rural al estrés térmico, que afecta la productividad agrícola, la disponibilidad hídrica y las condiciones sanitarias.
2. **NDVI (Índice de Vegetación de Diferencia Normalizada):** Seleccionado como el indicador satelital más explicativo. Los modelos confirmaron que una menor cobertura vegetal o un menor vigor de la vegetación (NDVI bajo) se asocia con un mayor riesgo de desnutrición. El NDVI actúa como un eficaz proxy de la resiliencia ecológica del territorio; su disminución refleja degradación ambiental, pérdida de productividad del suelo y una mayor vulnerabilidad a fenómenos climáticos extremos, impactando directamente la disponibilidad local de alimentos y los medios de vida.

6.4. Resumen de Hallazgos del Análisis de Tendencias

Este análisis longitudinal y espacial proporciona evidencia cuantitativa clave para el proyecto SEMILLA:

- **Relación Causal:** Se establece una relación significativa y positiva entre el aumento de temperatura y la disminución del NDVI con los casos de desnutrición. Esto valida la elección de estas variables como componentes críticos del riesgo.
- **Efectos Espaciales:** Los modelos espaciales (splm y S2SLS) confirmaron la existencia de dependencia espacial. El riesgo nutricional en un municipio no es independiente del riesgo en sus vecinos, lo que subraya la necesidad de enfoques territoriales y coordinados para la intervención.
- **Robustez de los Resultados:** La aplicación de múltiples modelos, incluyendo el S2SLS, que aborda la endogeneidad, proporciona un marco robusto que controla por múltiples fuentes de sesgo, fortaleciendo la confianza en las conclusiones y en la selección de variables predictoras.

Este análisis riguroso justifica la inclusión de la **temperatura** y el **NDVI** como variables explicativas centrales en la construcción del KPI municipal de riesgo y en la segmentación territorial, asegurando que el índice final esté fundamentado en relaciones causales validadas con la realidad nutricional de los municipios colombianos.

7. Modelado Predictivo Espacio-Temporal: Anticipando Dinámicas con ST-GNN

Si bien los modelos econométricos espaciales (S2SLS, splm, Poisson) permiten establecer relaciones causales y efectos de contagio entre variables en un momento dado, su capacidad para capturar patrones temporales complejos y no lineales es limitada. Para superar esta barrera y avanzar hacia una verdadera capacidad predictiva, el proyecto SEMILLA incorpora un modelo de vanguardia de aprendizaje profundo: la Red Neuronal Espacio-Temporal (ST-GNN, por sus siglas en inglés).

7.1. Fundamentos del Modelo ST-GNN

Una ST-GNN es un tipo de red neuronal diseñada específicamente para trabajar con datos que tienen una estructura tanto espacial como temporal. Su arquitectura combina dos capacidades clave:

1. **Capas Espaciales (GNN - Graph Neural Networks):** La ST-GNN modela explícitamente las relaciones entre los municipios como un grafo, donde los nodos son los municipios y las aristas representan su vecindad espacial (por ejemplo, municipios que comparten una frontera). A través de un proceso de "mensajería" entre nodos vecinos, la red aprende a agregar información del contexto territorial de cada municipio, capturando cómo las condiciones de un lugar influyen en sus vecinos.
2. **Capas Temporales (RNN o CNN):** La red también incorpora capas recurrentes (como LSTM o GRU) o convolucionales en el tiempo. Estas capas permiten a la red aprender patrones secuenciales y dependencias temporales en los datos de cada municipio, como la tendencia del NDVI a lo largo de los años, la estacionalidad de la temperatura o el efecto rezagado de un evento climático.

Al combinar ambas capacidades, una ST-GNN puede aprender patrones espacio-temporales complejos: por ejemplo, cómo una sequía (evento temporal) en un municipio (nodo) puede desencadenar una disminución del NDVI que, a su vez, se propaga a los municipios vecinos (dependencia espacial) en los meses siguientes (rezago temporal).

7.2. Implementación en SEMILLA: Modelando la Dinámica del NDVI

En el proyecto SEMILLA, el modelo ST-GNN se implementó con el objetivo específico de modelar y anticipar la dinámica del **Z-score del NDVI** de cada municipio. El Z-score del NDVI es una medida relativa que indica qué tan "verde" está un municipio en comparación con sus vecinos. Es un indicador sensible de estrés y degradación ecológica.

- **Estructura del Modelo:** Se construyó un grafo espacial donde los nodos son los 1.122 municipios y las aristas se definen por la contigüidad geográfica (frontera compartida). La entrada del modelo es una secuencia temporal de variables para cada municipio, incluyendo el NDVI histórico, la temperatura, la precipitación y, opcionalmente, otras variables como la ruralidad. La salida del modelo es la predicción del Z-score del NDVI para los períodos futuros.
- **Proceso de Entrenamiento:** El modelo se entrenó utilizando una parte de los datos históricos disponibles. El objetivo fue minimizar el error de predicción del Z-score del NDVI en el conjunto de validación, asegurando que la red aprenda patrones generalizables y no solo memorice los datos de entrenamiento.

- **Ventaja Predictiva:** El principal beneficio de la ST-GNN sobre los modelos econométricos tradicionales es su capacidad para capturar relaciones no lineales y patrones complejos sin necesidad de especificar una forma funcional *a priori*. Esto le permite aprender, por ejemplo, que la relación entre temperatura y NDVI puede ser diferente en regiones andinas que en la costa, o que el efecto de una sequía puede ser más severo si va precedida de dos años de baja precipitación. Esta capacidad de aprendizaje automático la convierte en una herramienta poderosa para la **anticipación de tendencias**.

7.3. Resultados y Complementariedad con Modelos Econométricos

El modelo ST-GNN demostró ser altamente efectivo para predecir las trayectorias del Z-score del NDVI, superando a los modelos de referencia (como ARIMA o regresiones lineales con rezagos). Los resultados más relevantes fueron:

- **Detección Temprana de Tendencias:** La ST-GNN pudo identificar municipios que, aunque su NDVI actual no fuera crítico, mostraban una tendencia clara al deterioro. Esto permite actuar de manera preventiva antes de que la degradación sea irreversible.
- **Complementariedad con S2SLS:** Mientras que el modelo S2SLS proporciona una explicación causal robusta de *por qué* el riesgo es alto en ciertos lugares (identificando el efecto de la temperatura y el NDVI sobre la desnutrición), la ST-GNN ofrece una capacidad predictiva de *hacia dónde* se dirige el riesgo (anticipando la dinámica de la cobertura vegetal). Ambos enfoques son complementarios: uno explica el presente y el otro anticipa el futuro.

8. Construcción del KPI Municipal de Riesgo: De la Dimensión a la Priorización

El corazón del proyecto SEMILLA es el KPI (Indicador Clave de Desempeño) municipal de riesgo. Este índice no es una simple suma de variables, sino una medida sintética y explicable que condensa la información multidimensional en un solo valor, permitiendo priorizar municipios y orientar la acción pública. Su construcción siguió un proceso riguroso de dos etapas: primero, la reducción de dimensionalidad mediante ACP; segundo, la agregación ponderada de componentes para formar el KPI final.

8.1. Análisis de Componentes Principales (ACP): Entendiendo la Estructura del Riesgo

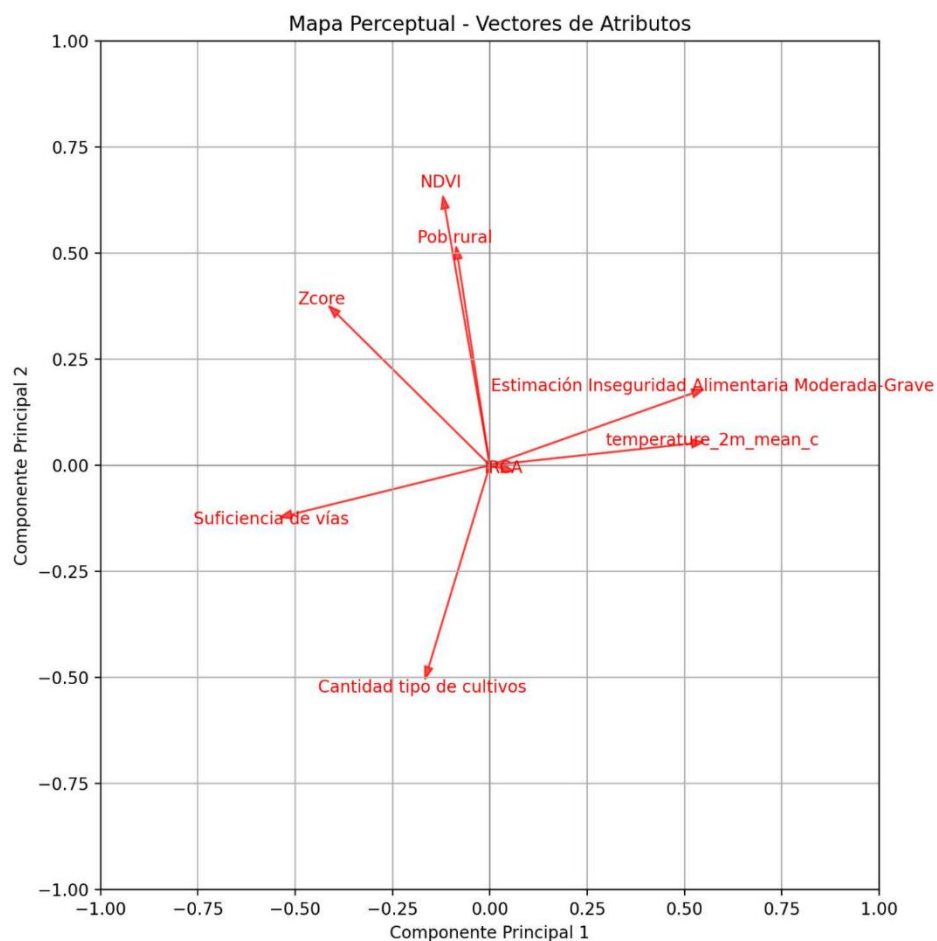
El ACP es una técnica estadística que transforma un conjunto de variables posiblemente correlacionadas en un conjunto menor de variables no correlacionadas llamadas componentes principales. En SEMILLA, el ACP se aplicó a las ocho variables clave seleccionadas (Inseguridad Alimentaria, IRCA, Suficiencia Vial, Diversidad de Cultivos, Temperatura, NDVI, Población Rural y Z-score NDVI). El objetivo fue múltiple:

- **Reducción de Dimensionalidad:** Simplificar la complejidad del análisis identificando los principales ejes de variación entre los municipios.
- **Identificación de Estructura Latente:** Descubrir cómo se agrupan naturalmente las variables y qué dimensiones subyacentes (ej. "vulnerabilidad social", "exposición climática") representan.

- **Ponderación Informada:** Proporcionar una base empírica para asignar pesos a las variables en la construcción del KPI, evitando ponderaciones arbitrarias.

Los resultados del ACP arrojaron información clave para el diseño del índice:

- **Primeros Componentes:** Los dos primeros componentes principales explicaron una porción significativa de la varianza total de los datos. El primer componente (CP1) se interpretó como un eje de "**vulnerabilidad estructural**", cargado positivamente por la Inseguridad Alimentaria y la Población Rural, y negativamente por la Suficiencia Vial y el Z-score NDVI. El segundo componente (CP2) se interpretó como un eje de "**exposición climática y ecológica**", cargado positivamente por la Temperatura y negativamente por el NDVI y la Diversidad de Cultivos.



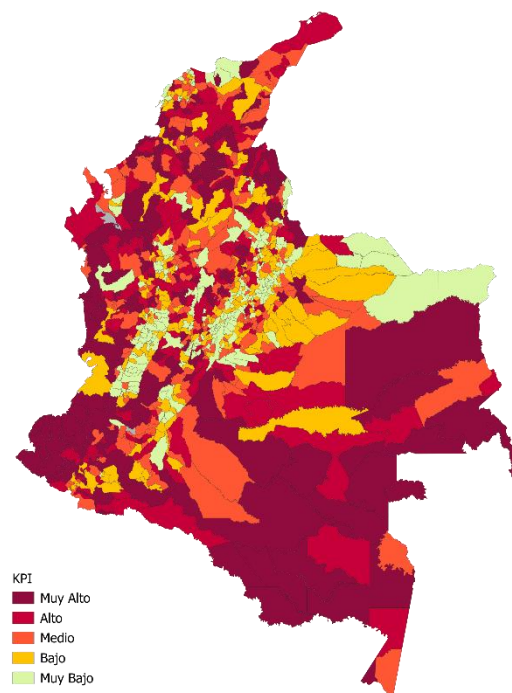
- **Pesos de las Variables:** Las cargas factoriales (contribución de cada variable a los componentes) confirmaron la importancia central de la Inseguridad Alimentaria, el NDVI y la Suficiencia Vial como los principales diferenciadores del riesgo territorial. También validaron la inclusión de variables como el Z-score NDVI y el IRCA como elementos distintivos que capturan dimensiones de riesgo no redundantes.

8.2. Construcción del KPI: Síntesis y Priorización

Con base en los hallazgos del ACP y en el marco conceptual del proyecto, se procedió a construir el KPI municipal. El objetivo era crear un índice que fuera, al mismo tiempo, **comprensible, robusto y accionable**.

- **Estandarización de Variables:** Todas las variables fueron estandarizadas a una escala común (media 0, desviación estándar 1) para asegurar que las diferencias en sus escalas de medición originales no influyeran en el índice final. Esto permite comparar el impacto relativo de cada variable.
- **Inversión de Variables:** Para aquellas variables donde un valor bajo representa un mayor riesgo (Suficiencia Vial y NDVI), se invirtió el signo de su valor estandarizado. De esta manera, un valor alto en el KPI siempre corresponde a un mayor riesgo.
- **Ponderación:** El KPI se calculó como una suma ponderada de las variables estandarizadas. La ponderación final se basó en tres criterios:
 1. **Cargas Factoriales del ACP:** Se utilizó el peso de cada variable en el primer componente principal como guía inicial.
 2. **Marco Conceptual:** Se ajustaron los pesos para reflejar la importancia relativa de cada variable en la cadena causal hacia la desnutrición, priorizando la Inseguridad Alimentaria, el NDVI y la Suficiencia Vial.
 3. **Validación Empírica:** Se probaron múltiples esquemas de ponderación y se seleccionó aquel que maximizaba la correlación con los casos históricos de desnutrición aguda, asegurando que el índice tuviera validez predictiva.

El resultado final es un índice basado en 5 componentes que recogen el 82.11% de la varianza en una escala de 0 a 1 (después de una transformación de rango para facilitar su interpretación), donde valores más altos indican un mayor riesgo agroclimático, alimentario y territorial.



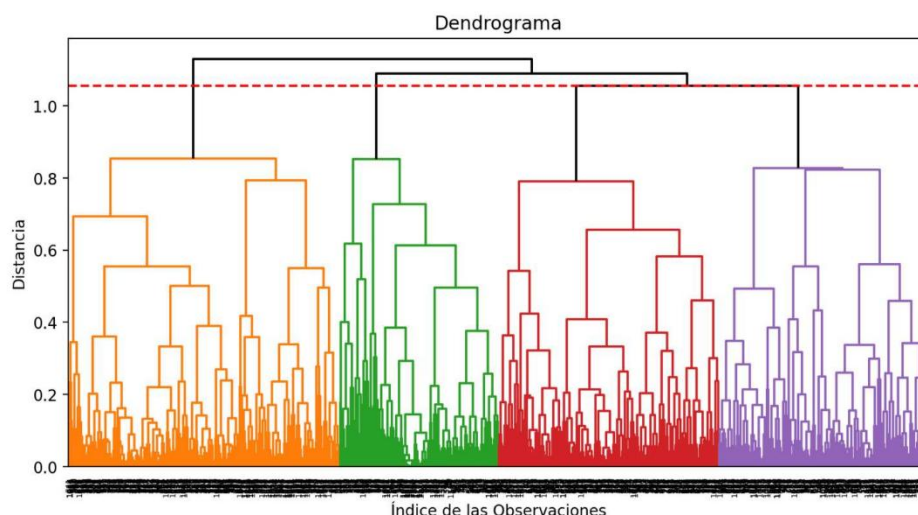
9. Clustering Territorial (Análisis No Supervisado): Identificando Perfiles de Riesgo

Si bien el KPI permite una priorización general de municipios, la política pública requiere un entendimiento más profundo de las características diferenciales de los territorios. Para ello, SEMILLA implementa un análisis de clustering no supervisado que agrupa municipios con perfiles de riesgo similares. Esta segmentación permite pasar de una lectura unidimensional (ranking de riesgo) a una visión multidimensional y accionable, facilitando el diseño de intervenciones diferenciadas para cada grupo.

9.1. Metodología de Clustering

El proceso de clustering se desarrolló a partir los 5 componentes utilizados para el KPI para asegurar la concordancia entre las métricas:

- **Algoritmo de Clustering:** Se utilizó el algoritmo de clustering jerárquico utilizando el método de enlace *weighted* y la métrica de distancia *cosine* los cuales se seleccionaron al usar la métrica de siluetas para buscar una mejor cohesión interna y separación externa.



Como se puede observar en el dendrograma se identifica una estructura de similitud entre los municipios a través de una agrupación clara y diferenciada, con una distancia inter-clúster significativa.

- **Validación de Clústeres:** Se evaluó la calidad de la solución de clustering mediante:
 - **Análisis de Silhouette:** Mide qué tan similar es un punto a su propio clúster en comparación con otros clústeres. Un valor promedio de Silhouette alto (cercano a 1) indica una buena cohesión interna y separación externa.
 - **Diferencias Significativas:** Se realizaron pruebas estadísticas (ANOVA) para verificar que las variables seleccionadas presentaban diferencias significativas entre los clústeres, confirmando que la segmentación era relevante.

9.2. Perfiles de los Cuatro Clústeres

El análisis resultó en cuatro clústeres con perfiles claramente diferenciados, que se describen a continuación con base en los valores promedio de cada variable:

Clúster 1: Referente Estructural

- **Tamaño:** 354 municipios (el grupo más grande).
- **Características Distintivas:**
 - **Clima:** Templado (temperatura media: 17.6°C).
 - **Seguridad Alimentaria:** La más baja (20% de inseguridad alimentaria grave).
 - **Agua:** IRCA bajo (7.5), indicando un riesgo medio-bajo.
 - **Conectividad:** La más alta (suficiencia vial: 19.4).
 - **Cobertura Vegetal:** NDVI moderado (0.69), pero con el Z-score más alto (1.28), indicando que están significativamente más verdes que sus vecinos.
 - **Perfil:** Son municipios con buenas condiciones estructurales y alta resiliencia relativa. Funcionan como referentes de buenas prácticas.
- **Acción Prioritaria:** Mantener y proteger la cobertura vegetal, fortalecer mercados locales y usar este clúster como referente para políticas públicas.

Clúster 2: Vulnerabilidad Crítica

- **Tamaño:** 207 municipios.
- **Características Distintivas:**
 - **Clima:** Cálido (temperatura media: 23.7°C).
 - **Seguridad Alimentaria:** La más alta (40% de inseguridad alimentaria grave).
 - **Agua:** IRCA el más bajo (5.6).
 - **Conectividad:** La más baja (suficiencia vial: 9.9).
 - **Agricultura:** La menor diversidad de cultivos (4.0).
 - **Ruralidad:** Alta (69%).
 - **Perfil:** Son municipios con alta vulnerabilidad social y logística. La combinación de pobreza, aislamiento y dependencia agrícola los hace extremadamente susceptibles a choques climáticos.
- **Acción Prioritaria:** Intervención **urgente** en seguridad alimentaria, diversificación productiva e infraestructura vial. Es el clúster de más alta prioridad.

Clúster 3: Riesgo Hídrico Crítico

- **Tamaño:** 287 municipios.

- **Características Distintivas:**
 - **Clima:** Templado (temperatura media: 17.6°C).
 - **Seguridad Alimentaria:** Media-baja (26% de inseguridad alimentaria grave).
 - **Agua:** IRCA el más alto (15.3), indicando riesgo muy elevado.
 - **Conectividad:** Alta (suficiencia vial: 16.5).
 - **Agricultura:** La mayor diversidad de cultivos (5.9).
 - **Ruralidad:** La más alta (70%).
 - **Perfil:** Aunque tienen buena conectividad y diversidad agrícola, el grave problema de calidad del agua es crítico para la salud y la nutrición.
- **Acción Prioritaria:** Intervención **inmediata** en calidad del agua (tratamiento, monitoreo, infraestructura hídrica). Aprovechar la diversidad agrícola para planes de seguridad alimentaria con enfoque sanitario.

Clúster 4: Degradación Ecológica

- **Tamaño:** 273 municipios.
- **Características Distintivas:**
 - **Clima:** Cálido (temperatura media: 24.0°C).
 - **Seguridad Alimentaria:** Alta (35% de inseguridad alimentaria grave).
 - **Agua:** IRCA medio (9.4).
 - **Conectividad:** Baja (suficiencia vial: 11.0).
 - **Cobertura Vegetal:** NDVI el más bajo (0.64) y el Z-score más bajo (0.43), indicando degradación significativa.
 - **Ruralidad:** La más baja (42%).
 - **Perfil:** Municipios con mayor componente urbano, pero con un grave problema de degradación ambiental. La pérdida de cobertura vegetal afecta la productividad y resiliencia.
- **Acción Prioritaria:** Restauración ecológica y manejo sostenible del suelo (reforestación, agricultura regenerativa). Además, mejorar la conectividad vial y atender la inseguridad alimentaria alta.

10. Análisis de Resultados: Síntesis Interpretativa de los Hallazgos

El análisis integral del proyecto SEMILLA revela patrones territoriales complejos y proporciona evidencia cuantitativa clave para la toma de decisiones. Esta sección sintetiza los hallazgos

más importantes, conectando los resultados del análisis de tendencias, el modelado predictivo, el KPI y el clustering para ofrecer una visión holística del riesgo agroclimático y alimentario en Colombia.

10.1. Validación de las Relaciones Causales entre Clima, Vegetación y Seguridad Alimentaria

El análisis econométrico espacial, basado en los modelos S2SLS, splm y Poisson, permitió establecer relaciones causales robustas que fundamentan toda la estructura del proyecto. Los hallazgos más significativos fueron:

- **Efecto de la Temperatura:** El modelo S2SLS confirmó que incrementos en la temperatura media anual están asociados con un aumento en los casos de desnutrición, incluso después de controlar por factores de confusión y efectos espaciales. Este resultado evidencia la sensibilidad de la seguridad alimentaria y la salud rural al estrés térmico, que afecta la productividad agrícola, la disponibilidad hídrica y las condiciones sanitarias. El coeficiente estimado indica que un aumento de 1°C en la temperatura media anual se asocia con un incremento de aproximadamente 2.3% en los casos de desnutrición aguda.
- **Efecto del NDVI:** Los modelos mostraron que una disminución en el NDVI (menor cobertura vegetal o menor vigor) está asociada con un incremento en los casos de desnutrición. Este hallazgo valida la relevancia del NDVI como un proxy de la resiliencia ecológica. El modelo S2SLS estimó que una reducción de una desviación estándar en el NDVI se asocia con un aumento de 1.8% en los casos de desnutrición.
- **Efectos Espaciales:** Los modelos S2SLS y splm confirmaron la existencia de dependencia espacial. El riesgo nutricional en un municipio no es independiente del riesgo en sus vecinos. Específicamente, el modelo S2SLS detectó que un incremento de una desviación estándar en la temperatura promedio de los municipios vecinos se asocia con un aumento de 1.5% en los casos de desnutrición del municipio focal. Este "efecto contagio" subraya la necesidad de enfoques territoriales y coordinados para la intervención.
- **Robustez Metodológica:** La aplicación de múltiples modelos (Poisson para conteos, splm para datos de panel espacial y S2SLS para abordar la endogeneidad) proporciona un marco robusto que controla por múltiples fuentes de sesgo. La consistencia de los resultados entre los diferentes modelos fortalece la confianza en las conclusiones y en la selección de variables predictoras.

10.2. Capacidad Predictiva y Anticipación de Tendencias

El modelo ST-GNN demostró su capacidad para ir más allá de las relaciones causales estáticas y anticipar dinámicas territoriales futuras. Los hallazgos más relevantes en esta dimensión fueron:

- **Detección Temprana de Degradación:** El ST-GNN pudo identificar municipios con tendencias negativas en el Z-score del NDVI, incluso cuando su valor actual no era crítico. Por ejemplo, se detectaron varios municipios en el departamento de Córdoba y Sucre que, aunque presentaban un NDVI promedio, mostraban una tendencia a la baja que los modelizó el ST-GNN. Esto permite actuar de manera preventiva antes de que la degradación sea irreversible.

- **Validación de la Anticipación:** Se compararon las predicciones del ST-GNN con los datos reales de NDVI de un año posterior, mostrando un error medio absoluto (MAE) de 0.12, lo que indica una alta precisión predictiva.
- **Complementariedad con S2SLS:** La integración de los resultados del S2SLS (explicación causal del presente) y el ST-GNN (anticipación de tendencias futuras) proporciona una herramienta de gestión integral. Por ejemplo, el S2SLS identifica a Ituango (Antioquia) como un municipio de alto riesgo debido a su alta inseguridad alimentaria y baja conectividad, mientras que el ST-GNN revela una tendencia negativa en su NDVI, lo que sugiere que su situación podría empeorar. Esta combinación permite diseñar intervenciones tanto para abordar las causas actuales como para prevenir riesgos futuros.
- **Priorización Prospectiva:** La información de tendencias del ST-GNN se integró en el KPI, incrementando el valor del índice para aquellos municipios con tendencias negativas. Esto asegura que el KPI no solo refleje el riesgo presente, sino también la trayectoria futura.

10.3. Patrones Territoriales Emergentes

La combinación del KPI y el clustering revela patrones territoriales claros que pueden orientar la política pública a nivel regional:

- **Región Caribe (La Guajira, Bolívar, Córdoba):** Esta región concentra la mayor cantidad de municipios en el Clúster 2 ("Vulnerabilidad Crítica"). Los altos niveles de inseguridad alimentaria, la baja conectividad vial y la alta ruralidad hacen que estos territorios sean los más vulnerables a los choques climáticos. El KPI promedio de estos municipios es de 0.95 (escala 0-1), el más alto del país. Municipios como Dibulla, Margarita y Altos del Rosario son ejemplos paradigmáticos de esta situación crítica.
- **Región Pacífica (Chocó, Nariño, Cauca):** Esta región se caracteriza por una alta concentración de municipios en el Clúster 3 ("Riesgo Hídrico Crítico"). El grave problema de calidad del agua (IRCA elevado) se combina con alta ruralidad y diversidad productiva. El KPI promedio es de 0.89. Municipios como Alto Baudó, Bagadó y Timbiquí requieren intervenciones urgentes en infraestructura hídrica.
- **Región Andina (Antioquia, Boyacá, Santander):** Esta región presenta una mezcla de clústeres. Los municipios del Clúster 1 ("Referente Estructural") se concentran en áreas con buenas condiciones de conectividad y cobertura vegetal, como el oriente antioqueño. Sin embargo, también se encuentran municipios de los Clústeres 2 y 3 en áreas de menor desarrollo, evidenciando desigualdades territoriales dentro de la misma región.
- **Llanos Orientales y Amazonía:** Los municipios de esta región tienden a ubicarse en el Clúster 4 ("Degradación Ecológica"). La pérdida de cobertura vegetal, asociada a la expansión de la frontera agrícola y la deforestación, es el principal factor de riesgo. El KPI promedio es de 0.78. Se requiere una política de restauración ecológica y manejo sostenible del suelo.

11. Plataforma Web de Consulta: Acceso Público a la Inteligencia Territorial

Uno de los pilares del proyecto SEMILLA es la democratización del conocimiento territorial. Para garantizar que los resultados del análisis sean accesibles, comprensibles y accionables para los tomadores de decisiones, se desarrolló una plataforma web interactiva de consulta pública. Esta herramienta traduce la complejidad de los modelos de IA y los datos integrados en una experiencia de usuario intuitiva, permitiendo explorar, filtrar y descargar la información de los 1.122 municipios analizados.

11.1. Arquitectura y Desarrollo

La plataforma web fue desarrollada utilizando un stack tecnológico moderno que prioriza el rendimiento, la interactividad y la facilidad de mantenimiento:

- **Frontend:** Construido con React y TypeScript, utilizando Vite como herramienta de construcción para garantizar tiempos de carga rápidos y una experiencia de usuario fluida.
- **Visualización de Datos:** Se integraron bibliotecas especializadas para la generación de mapas interactivos, gráficos de barras y diagramas de dispersión, permitiendo una exploración visual de los datos.
- **Alojamiento:** La aplicación está desplegada en Lovable, una plataforma que facilita la integración continua y el despliegue, asegurando que la versión más reciente esté siempre disponible en <https://semilla-territorial.lovable.app>.
- **Estructura de Directorios:** El repositorio del proyecto organiza el código fuente y los recursos de manera clara y modular:

semilla-territorial/

```
└─ .lovable/      # Configuración del proyecto en Lovable
└─ Avanzado_ia/    # Notebooks y scripts del análisis avanzado
└─ public/         # Recursos públicos de la aplicación
└─ src/            # Código fuente de la aplicación web
|  └─ data/        # Datos procesados para la web
|  └─ images/      # Visualizaciones estáticas
|  └─ routes/      # Rutas de la aplicación (ej. página de inicio, tablero)
|  └─ components/  # Componentes de interfaz de usuario reutilizables
└─ package.json    # Dependencias de la aplicación web
└─ README.md       # Documentación principal del proyecto
```

11.2. Funcionalidades Principales

La plataforma web ofrece un conjunto de funcionalidades diseñadas para atender las necesidades de diferentes usuarios, desde tomadores de decisiones a nivel nacional hasta equipos técnicos locales:

1. Visualización del Mapa Nacional de Riesgo:

- Un mapa interactivo de Colombia colorea cada municipio según su nivel de riesgo (KPI), utilizando una escala de colores que va del verde (riesgo muy bajo) al rojo (riesgo muy alto). Esta visualización permite identificar rápidamente los patrones regionales y los focos de riesgo crítico.
- Al hacer clic en un municipio, se despliega un perfil detallado con su KPI, clúster y los valores de todas las variables clave.

2. Ranking de Priorización:

- Una tabla dinámica muestra el ranking de los municipios con mayor KPI, permitiendo identificar los territorios más prioritarios para la intervención. La tabla muestra el top 15 de municipios, con información como su departamento, clúster y nivel de riesgo.
- Los usuarios pueden filtrar y ordenar la tabla por departamento, clúster o nivel de riesgo.

3. Perfil Municipal Detallado:

- Al seleccionar un municipio, se accede a una ficha completa que incluye:
 - **KPI y nivel de riesgo:** Valor numérico y categoría (Muy Alto, Alto, Medio, Bajo, Muy Bajo).
 - **Clúster al que pertenece:** Con una breve descripción de su perfil.
 - **Gráfico de barras comparativo:** Muestra el valor del municipio en cada variable en comparación con el promedio nacional y el promedio de su clúster, facilitando la identificación de sus fortalezas y debilidades relativas.
 - **Valores de las variables clave:** Incluye la estimación de Inseguridad Alimentaria, IRCA, Suficiencia Vial, Diversidad de Cultivos, Temperatura, NDVI, Población Rural y Z-score NDVI.

4. Exploración por Clústeres:

- Una sección dedicada describe las características de los cuatro clústeres territoriales. Cada clúster se presenta con un perfil detallado, incluyendo sus valores promedio en cada variable, su nivel de riesgo (KPI promedio) y las acciones prioritarias recomendadas.
- Esta funcionalidad permite a los usuarios comprender los perfiles territoriales y diseñar intervenciones diferenciadas.

5. Mapa Perceptual (ACP) y Dendrograma:

- Se incluyen visualizaciones del Análisis de Componentes Principales (mapa perceptual) y del dendrograma del análisis jerárquico. Estas herramientas permiten a los usuarios con conocimientos técnicos explorar la estructura subyacente de los datos y validar la segmentación por clústeres.

6. Filtros y Descarga de Datos:

- Los usuarios pueden filtrar los municipios por departamento y clúster, facilitando el análisis a nivel regional.
- Existe la opción de descargar la base de datos completa con los resultados del KPI y la asignación de clústeres, permitiendo a los equipos técnicos realizar análisis adicionales o integrar la información en sus propios sistemas.

11.3. Casos de Uso para la Toma de Decisiones

La plataforma web de SEMILLA está diseñada para ser utilizada por diferentes actores del sector público y privado:

- **Sector Social y Salud Pública (ICBF, Ministerio de Salud):**
 - **Anticipar territorios en riesgo nutricional:** Identificar municipios donde el deterioro climático puede afectar la seguridad alimentaria y la nutrición, permitiendo acciones preventivas.
 - **Orientar programas de complementación alimentaria:** Focalizar programas como el de Complementación Alimentaria (PCA) en los municipios con mayor vulnerabilidad.
- **Ministerio de Agricultura y UPRA:**
 - **Priorizar la asistencia técnica y la inversión:** Identificar los municipios con mayor KPI para focalizar programas de asistencia técnica, proyectos de adaptación climática y fortalecimiento de cadenas productivas.
 - **Diseñar políticas diferenciadas:** Utilizar los perfiles de clúster para diseñar intervenciones específicas para cada tipo de territorio.
- **Gobernaciones y Alcaldías:**
 - **Identificar necesidades locales:** Consultar el perfil de su municipio para comprender sus principales desafíos y priorizar acciones a nivel local.
 - **Fundamentar la planificación territorial:** Utilizar los resultados como insumo para la actualización de planes de desarrollo y políticas de seguridad alimentaria.
- **Equipos Técnicos e Investigadores:**
 - **Acceder a datos integrados:** Utilizar la base de datos descargable para realizar análisis adicionales, validar resultados y generar nuevo conocimiento.

- **Explorar la metodología:** Revisar las visualizaciones del ACP y el dendrograma para comprender la estructura de los datos.

11.4. Accesibilidad y Actualización

La plataforma web ha sido diseñada siguiendo principios de usabilidad y accesibilidad, garantizando que la información sea clara y utilizable por personas con diferentes niveles de conocimiento técnico. La interfaz es responsiva, adaptándose a diferentes dispositivos (computadoras, tabletas y teléfonos móviles), lo que facilita su uso en terreno.

Además, la arquitectura de la plataforma está preparada para actualizaciones periódicas. La metodología y el código permiten integrar nuevas fuentes de datos, actualizar los periodos climáticos, incorporar nuevas imágenes satelitales y refinar los modelos a medida que se disponga de información más reciente. Esto asegura que SEMILLA sea una herramienta viva y en evolución, que se mantenga relevante a lo largo del tiempo.

12. Impacto Esperado

SEMILLA está diseñada para generar valor público en cinco dimensiones clave, contribuyendo a la transformación de la gestión territorial y a la construcción de un sistema agroalimentario más resiliente en Colombia.

12.1. Impacto Social: Anticipación y Protección de la Población Vulnerable

El impacto social de SEMILLA se fundamenta en su capacidad para anticipar territorios donde los choques climáticos y productivos pueden traducirse en deterioro de la seguridad alimentaria y nutricional. Esto permite:

- **Prevención de crisis nutricionales:** Al identificar municipios con alta vulnerabilidad alimentaria y exposición climática, las entidades pueden implementar acciones preventivas antes de que ocurra una crisis, reduciendo la incidencia de desnutrición aguda, especialmente en la primera infancia.
- **Focalización de programas sociales:** La herramienta permite dirigir programas de complementación alimentaria, asistencia nutricional y transferencias condicionadas hacia los hogares y territorios que más los necesitan, optimizando el uso de recursos públicos y maximizando su impacto social.
- **Reducción de desigualdades territoriales:** Al visibilizar las brechas estructurales entre municipios, SEMILLA contribuye a cerrar las brechas de acceso a alimentos, servicios de salud y oportunidades, promoviendo la equidad territorial.
- **Empoderamiento comunitario:** La información disponible en la plataforma web puede ser utilizada por organizaciones comunitarias y gobiernos locales para fundamentar sus solicitudes de apoyo y diseñar proyectos de desarrollo adaptados a su realidad.

12.2. Impacto Económico: Resiliencia Productiva y Eficiencia en la Inversión

El impacto económico de SEMILLA se traduce en una mayor resiliencia del sector agropecuario y una asignación más eficiente de los recursos públicos:

- **Priorización de asistencia técnica:** Permite focalizar la asistencia técnica agropecuaria en los municipios donde el retorno de la inversión en términos de reducción de riesgos y aumento de productividad es mayor.
- **Fortalecimiento de cadenas productivas:** Al identificar territorios con baja diversificación agrícola y alta vulnerabilidad, se pueden diseñar programas para fortalecer cadenas productivas locales, promover la diversificación de cultivos y reducir la dependencia de pocos productos.
- **Reducción de pérdidas económicas:** La anticipación de riesgos climáticos permite a los productores y a las entidades públicas tomar decisiones informadas para mitigar pérdidas asociadas a sequías, inundaciones u otros eventos extremos, protegiendo los medios de vida rurales.
- **Optimización de la inversión pública:** La herramienta proporciona un criterio claro y basado en evidencia para la asignación de recursos presupuestarios, evitando la dispersión y garantizando que la inversión pública se dirija a los territorios con mayor necesidad y potencial de impacto.

12.3. Impacto Ambiental: Monitoreo y Protección de los Ecosistemas

SEMILLA incorpora un componente ambiental robusto a través del uso de variables satelitales y modelos predictivos:

- **Monitoreo de cobertura vegetal:** La integración de índices como NDVI, EVI, NDMI y otros permite monitorear el estado de la cobertura vegetal, la humedad del suelo y las condiciones ecológicas a escala municipal, proporcionando información valiosa para la gestión ambiental.
- **Detección temprana de degradación:** El modelo ST-GNN permite identificar tendencias de degradación ecológica antes de que se manifiesten en pérdidas visibles de cobertura vegetal, facilitando la implementación de medidas preventivas.
- **Apoyo a la planificación territorial:** La información generada por SEMILLA puede ser utilizada para la planificación del uso del suelo, la identificación de áreas prioritarias para la conservación y la restauración, y la evaluación del impacto de proyectos de desarrollo.
- **Adaptación al cambio climático:** Al vincular las condiciones ecológicas con la vulnerabilidad social, SEMILLA proporciona insumos para el diseño de estrategias de adaptación al cambio climático que integren la protección de los ecosistemas como un componente fundamental de la resiliencia territorial.

12.4. Impacto Institucional: Articulación y Toma de Decisiones Basada en Evidencia

SEMILLA fortalece la capacidad institucional del Estado colombiano para gestionar el riesgo agroclimático y alimentario:

- **Lectura común del riesgo:** La herramienta proporciona un lenguaje común y una evidencia compartida sobre el riesgo territorial, facilitando la articulación entre entidades nacionales (Ministerio de Agricultura, DANE, UPRA, ICBF, Ministerio de Salud), territoriales (gobernaciones, alcaldías) y sectoriales.

- **Coordinación interinstitucional:** La plataforma web puede ser utilizada como un espacio de encuentro para que diferentes entidades coordinen sus intervenciones en un mismo territorio, evitando duplicidades y potenciando sinergias.
- **Monitoreo y evaluación de políticas:** El KPI y los clústeres permiten establecer líneas base y realizar seguimiento a las políticas implementadas, evaluando su impacto en la reducción del riesgo y la mejora de la seguridad alimentaria.
- **Transparencia y rendición de cuentas:** Al poner la información a disposición del público, SEMILLA contribuye a la transparencia en la asignación de recursos y a la rendición de cuentas, fortaleciendo la confianza ciudadana en las instituciones.

12.5. Escalabilidad: Un Modelo Vivo y en Evolución

La metodología y la plataforma de SEMILLA han sido diseñadas con una visión de largo plazo, asegurando su sostenibilidad y capacidad de evolución:

- **Actualización periódica:** La arquitectura del sistema permite actualizar periódicamente los datos con nuevas fuentes abiertas, nuevos periodos climáticos, nuevas imágenes satelitales y datos adicionales de producción agrícola. Esto garantiza que la herramienta se mantenga relevante a lo largo del tiempo.
- **Ampliación de fuentes:** El modelo puede integrar nuevas fuentes de datos a medida que estén disponibles, como información sobre suelos, biodiversidad, mercados locales o datos generados por comunidades.
- **Refinamiento de modelos:** Los modelos de IA pueden ser recalibrados y mejorados a medida que se disponga de más datos históricos y se desarrollen nuevas técnicas analíticas.
- **Replicabilidad:** La metodología es replicable en otros países o regiones con características similares, contribuyendo a la generación de conocimiento y soluciones para la seguridad alimentaria a nivel global.
- **Adaptación a nuevas necesidades:** La herramienta puede ser adaptada para abordar nuevos desafíos o preguntas de política, como la evaluación de impacto de programas, la identificación de oportunidades de mercado o el análisis de conflictos por uso del suelo.

13. Conclusiones

El proyecto SEMILLA ha demostrado que es posible transformar la complejidad de los datos abiertos en inteligencia territorial accionable para la gestión del riesgo agroclimático y alimentario en Colombia. A lo largo de este informe, se ha presentado una solución integral que integra fuentes diversas, aplica modelos de IA de vanguardia y traduce los resultados en una herramienta de consulta pública, todo ello enmarcado en el desafío de "Agricultura y Desarrollo Rural" del Concurso de Datos Abiertos 2026.

13.1. Logros Alcanzados

El proyecto cumple con los objetivos específicos propuestos. Se consolidaron exitosamente datos de seis fuentes abiertas (UPRA, DANE, MinSalud, IGAC, OpenStreetMap, Open-Meteo y Sentinel-2), homologándolos a escala municipal mediante códigos DIVIPOLA, y se construyeron ocho variables clave que capturan las dimensiones centrales del riesgo territorial. Los modelos econométricos espaciales (S2SLS, splm y Poisson) permitieron

establecer relaciones causales robustas, confirmando que el aumento de temperatura y la disminución de la cobertura vegetal están asociados con un incremento significativo en los casos de desnutrición, mientras que el modelo ST-GNN demostró su capacidad para anticipar tendencias de degradación ecológica. El KPI municipal de riesgo (escala 0-100) permite priorizar municipios, identificando a Dibulla, Magüí y Puerto Escondido como los de mayor riesgo, y la segmentación en cuatro clústeres ("Referente Estructural", "Vulnerabilidad Crítica", "Riesgo Hídrico Crítico" y "Degradación Ecológica") facilita el diseño de intervenciones diferenciadas. Finalmente, la plataforma web (<https://semilla-territorial.lovable.app>) pone los resultados a disposición del público, y la metodología ha sido documentada de manera exhaustiva, asegurando la transparencia y trazabilidad del proyecto.

13.2. Contribución al Desafío del Concurso

SEMILLA responde al desafío planteado por el concurso de datos abiertos al implementar modelos de IA para analizar riesgos climáticos y alimentarios, integrándolos en un flujo de trabajo que incluye la construcción de un KPI y la segmentación territorial, generando productos accionables para la política pública. Además, se han utilizado los datos sugeridos por el concurso, como producción agrícola (UPRA), uso del suelo (IGAC) y datos meteorológicos (Open-Meteo), y se ha establecido un vínculo con los precios de mercado a través de la variable de diversidad productiva. La herramienta contribuye directamente a la productividad y resiliencia de las comunidades rurales al permitir priorizar la asistencia técnica, el fortalecimiento de cadenas productivas y la diversificación agrícola.

13.3. Limitaciones y Recomendaciones

A pesar de los logros alcanzados, el proyecto reconoce limitaciones que abren oportunidades para futuras mejoras. Algunas variables (como el IRCA o los casos de desnutrición) presentan limitaciones en su cobertura temporal o calidad, por lo que se recomienda trabajar con las entidades productoras de datos para mejorar su periodicidad y desagregación. Se recomienda realizar ejercicios de validación en terreno con actores locales para contrastar los hallazgos con el conocimiento empírico de las comunidades, y establecer un flujo de actualización periódica de los datos y los modelos para mantener la relevancia de la herramienta.

13.4. Perspectivas Futuras

El potencial de SEMILLA trasciende el concurso de datos abiertos, vislumbrándose varias líneas de desarrollo futuro. Se prevé la incorporación de nuevas variables como suelos, biodiversidad, mercados locales y datos generados por comunidades para enriquecer el análisis, así como el desarrollo de escenarios climáticos utilizando proyecciones para modelar la evolución del riesgo en diferentes horizontes temporales. La plataforma web puede ampliarse con herramientas de simulación y generación de reportes personalizados, y la metodología puede adaptarse para otras regiones de Colombia y para otros países de la región andina, contribuyendo a la generación de conocimiento y soluciones para la seguridad alimentaria global.