



ARQUITECTURA ORIENTADA A OBJETOS PARA ANÁLISIS DE DATOS EN AGRICULTURA DE PRECISIÓN

Alumno: Matías Massigoge N°2750/5.

Director: Ph. D. Silvia Gordillo.

Codirector: Ph. D. Rodolfo Bongiovanni.

Noviembre de 2006

TES 06/10 DIF-02923 SALA	 <p>UNIVERSIDAD NACIONAL DE LA PLATA FACULTAD DE INFORMATICA Biblioteca 50 y 120 La Plata catalogo.info.unip.edu.ar biblioteca@info.unip.edu.ar</p>  <p>DIF-02923</p>
---	---

Índice General

1. Introducción	1
1.1. Avances tecnológicos	1
1.2. Agricultura de Precisión	1
1.3. Complejidad de los sistemas de información geográfica	2
1.4. Beneficios al nuevo escenario de la producción agraria	2
1.5. Propuesta de la tesis	3
1.5.1. Abstracción	3
1.5.2. Encapsulamiento	3
1.5.3. Herencia	3
1.5.4. Reusabilidad	3
1.5.5. Extensibilidad	3
1.6. Estructura de la tesis	3
2. Revisión Bibliográfica	5
2.1. Agricultura de Precisión (AP)	5
2.2. Aplicaciones de la Agricultura de Precisión	5
2.3. Herramientas y tecnologías involucradas en la Agricultura de Precisión	6
2.3.1. Sistemas de Posicionamiento Global (GPS)	6
2.3.2. GPS Diferencial (DGPS)	7
2.3.3. Monitor de rendimiento	7
2.3.4. Muestreo de Suelos	8
2.3.5. Sensores remotos	8
2.3.6. Sistemas de Información Geográfica (GIS)	9
2.3.7. Tratamientos con Dosificación Variable (VRT)	9
2.4. Aspectos que deben tenerse en cuenta con la Agricultura de Precisión	10
2.5. Análisis de Datos en la AP	11
2.6. Diferentes análisis de datos que se llevan a cabo en la AP	11
2.7. Metodologías y herramientas utilizadas en el análisis de la AP	12
2.8. Soluciones de Software existentes	12
2.9. Problemas de dominio de GIS	13
2.10. Tecnología Orientada a Objetos (OO)	14
2.11. Beneficios de la Tecnología Orientada a Objetos	14

2.12.	Problemas con las soluciones OO	15
2.13.	Unified Modeling Language (UML)	15
2.14.	Patrones de Diseño OO	15
3.	Etapas y características de un proceso de análisis de datos de un monitor de rendimiento.	17
3.1.	Filtrado de Datos	17
3.2.	Modelado de Datos	18
3.3.	Definición de Zonas de Manejo	19
3.4.	Agregación de información sobre tratamientos	19
3.5.	Cálculos adicionales con planillas de cálculo	20
3.6.	Análisis estadísticos	20
3.6.1.	Problemas con análisis estadísticos tradicionales	20
3.6.2.	Técnicas de regresión espacial en el análisis estadístico	20
3.6.3.	Econometría espacial	21
3.7.	Interpretación de los resultados	23
4.	Arquitectura de la estrategia de análisis de datos	25
4.1.	Arquitectura del sistema	25
4.2.	Modelado y análisis de datos	25
4.2.1.	El Lote	26
4.2.2.	Las Campañas	27
4.2.3.	Estrategias de Análisis:	27
4.2.4.	Los Procesos	27
4.3.	Análisis de Casos de Uso	28
4.4.	Metodología	29
4.5.	Resultado	29
4.5.1.	Modelo para incorporación de características geográficas a elementos del dominio	29
4.5.2.	El Lote y las Campañas	30
4.5.3.	Estrategias de Análisis:	30
4.5.4.	Los Procesos	31
4.5.4.1.	Filtrado de Datos	32
4.5.4.2.	Modelado	32
4.5.4.3.	Determinación de Zonas de Manejo	33

4.5.4.4.	Análisis Estadístico	33
4.5.4.5.	Análisis Económico	34
4.5.4.6.	Generación de Recomendaciones	35
4.6.	Estrategia completa de análisis	36
4.7.	Arquitectura completa:	37
5.	Instancia de la arquitectura	39
5.1.	El Lote y las Campañas	39
5.2.	Estrategias de Análisis:	39
5.3.	Los Procesos	40
5.3.1.	Filtrado de Datos	40
5.3.2.	Modelado	41
5.3.3.	Determinación de Zonas de Manejo	41
5.3.4.	Análisis Estadístico	42
5.3.5.	Análisis Económico	43
5.3.6.	Generación de Recomendaciones	43
5.4.	Arquitectura completa	45
5.5.	Diagrama de secuencia del análisis utilizando econometría espacial	46
6.	Conclusiones	47
6.1.	Conclusiones	47
6.2.	Trabajos futuros	48
7.	Referencias Bibliográficas	50

Capítulo 1: Introducción

La actividad agraria que se lleva a cabo en la Argentina sobre grandes extensiones de tierra hace uso intensivo de herramientas tecnológicamente complejas.

Tradicionalmente los productores agrarios argentinos han aceptado la incorporación de nuevas tecnologías en sus actividades.

La utilización de estas herramientas con alta incorporación de tecnología y de alto costo genera una necesidad de gestión sobre los recursos, tanto para tomar decisiones sobre su adopción como para su posterior implementación en las actividades: la adopción es deseable llevarla a cabo sólo si se prevé recuperar la inversión. La implementación requiere de análisis complejos para que sea exitosa, ya que se utilizan a lo largo de una campaña diversos insumos como semillas, fertilizantes y herbicidas en cantidades que es necesario cuidar para lograr una maximización de la rentabilidad de la inversión y proteger el recurso principal que es la tierra.

1.1. Avances tecnológicos

Los últimos avances incluyen técnicas que permiten manejar de manera diferenciada de acuerdo a la posición geográfica la aplicación de insumos en función de las necesidades específicas del sitio. Tradicionalmente la cantidad de estos insumos que se aplicaban a los lotes eran uniformes y estaban indicadas por recomendaciones agronómicas para la zona en cuestión, como consecuencia de este manejo se producían tanto subaplicaciones en unos sectores como sobreaplicaciones en otros por pertenecer al mismo lote con la consiguiente merma en la rentabilidad y el desperdicio de insumos.

1.2. Agricultura de Precisión

La actividad con manejo diferenciado es llamada Agricultura de Precisión (AP), también conocida como manejo sitio-específico, donde es indispensable recolectar grandes volúmenes de información para su posterior análisis y generación de recomendaciones.

Tanto las tareas de almacenamiento como su posterior análisis requieren la utilización de sistemas complejos conocidos como Sistemas de Información Geográfica (GIS).

1.3. Complejidad de los sistemas de información geográfica

Sin embargo, trabajar con productos orientados a GIS no es sencillo y en muchas situaciones puede resultar frustrante:

- Con excepción de algunos productos específicos, los GIS existentes son de uso general y con interfaces que dificultan su utilización para usuarios no expertos.
- No poseen herramientas de análisis estadístico específicas para agricultura volviendo necesaria la interacción con otros sistemas en varios pasos del análisis.
- Los formatos de datos de muestras que generan la utilización de herramientas agropecuarias no son estándar por lo tanto es necesario llevar a cabo procesos de transformación a los mismos previamente al análisis y en la generación de recomendaciones.

1.4. Beneficios al nuevo escenario de la producción agraria

Hasta principios de la década del 1990, la producción básicamente la realizaban productores dueños de la tierra o productores capitalizados en maquinaria que además de trabajar su tierra tenían arrendados lotes de vecinos.

A partir de mediados de la década del 1990, muchos productores dejaron su espacio a otros tipo de organizaciones más dinámicas y más flexibles (como pools de siembra y contratistas altamente capitalizados), cuyas decisiones se toman principalmente en la rentabilidad del negocio, por lo que la producción se hizo necesariamente más eficiente y cada día más demandante de información, tanto de mercados como agronómica.

Hoy en día muchos grupos de productores (tanto terratenientes como arrendatarios) le agregan valor a la superficie trabajada a través de la acumulación de información geográficamente referenciada, y auguran que en un futuro próximo se valorizara más ese cúmulo de información que la tierra en sí.

Para este escenario se vuelve fundamental el plantear la gestión de cada campaña como un sistema de información donde no sólo se puedan buscar respuestas para el momento sino que permitan disponer de los datos para futuros análisis.

1.5. Propuesta de la tesis

En esta Tesis se propone una arquitectura Orientada a Objetos (OO) que surge de un análisis exhaustivo a los pasos comunes a los diferentes análisis revisados, permitiendo modelar cada estrategia de análisis como una composición de pasos, beneficiándose con las características inherentes a la Orientación a Objetos:

- 1.5.1. Abstracción: foco en las características esenciales e inherentes a una entidad
- 1.5.2. Encapsulamiento: separación de los aspectos externos de un objeto de su implementación interna.
- 1.5.3. Herencia: los objetos se relacionan en una estructura jerárquica en el que los hijos heredan atributos y comportamiento de sus padres.
- 1.5.4. Reusabilidad: la abstracción, el encapsulamiento y la herencia favorece la reutilización objetos previamente definidos en nuevas aplicaciones.
- 1.5.5. Extensibilidad: todas estas características generan un ambiente en el que los componentes del sistema pueden ser reemplazados o combinados de nuevas maneras para cubrir nuevos requerimientos.

Esta arquitectura puede llevarse a una instancia que contemple los pasos de un análisis en particular, solucionando todo el proceso durante el cual es posible interactuar con otras herramientas para resolver procesos específicos.

1.6. Estructura de la tesis

El **Capítulo 2** revisa la bibliografía referente a Agricultura de Precisión, sus diferentes estrategias de análisis y las características que hay que tener en cuenta al elegir una en particular. También se cubren los aspectos tecnológicos y se presenta el paradigma que se utilizará en la búsqueda de la solución desde las herramientas informáticas.

El **Capítulo 3** profundiza el análisis de los datos que participan en la estrategia, tanto su adquisición como sus transformaciones en resultados intermedios y finales teniendo en cuenta los pormenores que deben tenerse en cuenta en cada una de estas etapas para llegar a conclusiones confiables. En cada uno de los análisis cubiertos se propone en forma de conclusiones las características fundamentales que serán tenidas en cuenta en la generación del modelo propuesto.

El **Capítulo 4** propone la arquitectura abstracta que resuelve el problema de manera general y que servirá de base para los casos particulares que puedan surgir en la actividad.

El **Capítulo 5** muestra una instancia de la arquitectura propuesta para resolver la estrategia llevada a cabo por el Ph. D. Rodolfo Bongiovanni en su tesis doctoral que involucra técnicas complejas de estadística para contemplar la correlación espacial de las muestras llamada Econometría Espacial, utilizada para estimar la respuesta del rendimiento a la aplicación de Nitrógeno y análisis económicos para estimar la rentabilidad del tratamiento.

Finalmente el **Capítulo 6** presenta las Conclusiones obtenidas a partir del desarrollo de esta Tesis y en el **Capítulo 7** se enumera la bibliografía utilizada a lo largo de la misma.

Capítulo 2: Revisión Bibliográfica

En el presente capítulo se recorrerán diversas fuentes bibliográficas sobre el tema de interés. Primero se revisaran los temas correspondientes a la Agricultura de Precisión y luego se analizarán los diferentes paradigmas que se utilizarán en el desarrollo de la solución.

2.1. Agricultura de Precisión (AP)

Tradicionalmente en las prácticas agronómicas se asume que las características del suelo y su respuesta a la siembra y a los distintos tratamientos dentro de un lote son uniformes. Se aplica la misma cantidad de semillas, fertilizantes, herbicidas, riego, etc. a toda la superficie en base a recomendaciones agronómicas para la zona en cuestión.

La Agricultura de Precisión es el uso de la tecnología de la información, tales como Sistemas de Información Geográfica (SIG), Sistemas de Posicionamiento Global (GPS), etc. para adecuar el manejo de suelos y cultivos a la variabilidad presente dentro de un lote (Lowenberg-DeBoer and Swinton, 1997). Permite tratar áreas menores dentro del lote de manera diferente a como se trataría el lote entero (Roberts, 2000), esto significa adaptar la técnica de cultivo a las condiciones particulares que se presentan en cada uno de los sectores del campo (Baños y Goenaga, 2003).

Este manejo diferencial traería aparejado la disminución de la contaminación, el aumento de la producción y un mejor resultado económico, ya que las recomendaciones tradicionales que se basan en el promedio del lote provocan una sobreaplicación en sectores con el consiguiente riesgo ambiental y una subaplicación en otros provocando una pérdida de rendimiento (Bongiovanni y Lowenberg-DeBoer, 2006).

2.2. Aplicaciones de la Agricultura de Precisión

Desde el punto de vista de la empresa agropecuaria, la AP cubre tres áreas principales: (1) Producción, (2) Administración, y (3) Medio Ambiente.

Desde el punto de vista (1) microeconómico o de la producción, las técnicas de la AP pueden ser aplicadas, por ejemplo, para la búsqueda de (a) la eficiencia técnica y económica

(rentabilidad), (b) el control de la calidad, (c) la trazabilidad de la producción, (d) el control de plagas y enfermedades, (e) la fertilización con dosis variable (DV), etc.

La AP también puede ser adoptada como (2) una herramienta de gestión, para: (a) registrar datos, (b) supervisar empleados, (c) controlar actividades, (d) controlar a los contratistas y (e) capacitar a los operarios.

Por último, pero no menos importante, (3) la AP es una herramienta útil para proteger el medio ambiente y contribuir a la sustentabilidad de la agricultura (Bongiovanni y Lowenberg-DeBoer, 2006)

2.3. Herramientas y tecnologías involucradas en la Agricultura de Precisión

Existe un conjunto de herramientas que se utilizan dentro de la actividad, muchas de ellas de reciente aparición en el mercado.

2.3.1. Sistemas de Posicionamiento Global (GPS)



Imagen 1: Dispositivos de Posicionamiento Global

Fueron desarrollados por el Departamento de Defensa de los Estados Unidos con fines militares pero hoy día su aplicación en diferentes campos sobrepasa su aplicación en aspectos militares. Se trata de un sistema de 24 satélites que orbitan la tierra. A través de sus señales, es posible determinar la posición (longitud y latitud) de objetos en la superficie terrestre. Es una potente herramienta para ser utilizada en la AP con el fin localizar un punto exacto donde se han tomado muestras para medir variabilidad espacial en una serie temporal sobre aspectos tales como fertilidad del suelo, rendimiento de cultivos, densidades de plagas, etc. (Emmen, 2004)

2.3.2. GPS Diferencial (DGPS)

El sistema DGPS usa un monitor receptor estacionario para calcular la diferencia entre la posición real de un punto la determinada por el sistema de satélites en un momento dado. Esto permite que el error inducido como parte de la desclasificación militar pueda ser eliminado. Con esto se dispone de una precisión en tres dimensiones de 20 centímetros o aún menor (Goddard et al., 1995).

2.3.3. Monitor de rendimiento

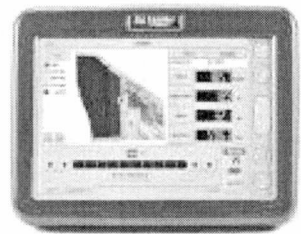
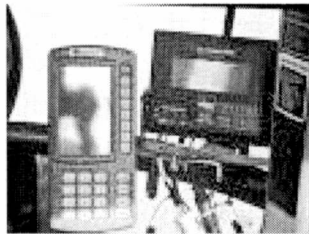
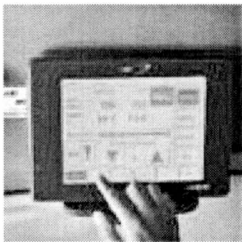


Imagen 2: Monitores de Rendimiento

El monitor es esencialmente una pequeña computadora que procesa y registra en la cosechadora la información de rendimiento, almacenándola en su memoria. Luego, mediante una tarjeta de memoria se transfiere la información a una computadora de escritorio. Los datos necesarios son relevados mediante sensores montados en la cosechadora, los cuales reportan a un monitor conectado al DGPS (Baños y Goenaga, 2003).

Los datos resultan valiosos ya que pueden cuantificarse mediante el uso de programas de computación específicos obteniendo información útil para aplicar en futuras campañas.

El mapa de rendimiento es la representación gráfica del rendimiento en grano del cultivo. (von Martini et al., 2003).

2.3.4. Muestreo de Suelos



Imagen 3: Actividades de muestreo de suelo

Los muestreos intensivos de suelo permiten evaluar las características físico - químicas (nivel de nutrientes, pH, etc.) de cada sector del lote. Las muestras se pueden extraer de grillas rectangulares delimitadas en el terreno y ubicadas mediante DGPS (usualmente cuadrículas de una hectárea) o de sectores del lote con características homogéneas (por tipo de suelos o por zonas de manejo), también ubicándolas mediante DGPS (Baños y Goenaga, 2003).

2.3.5. Sensores remotos

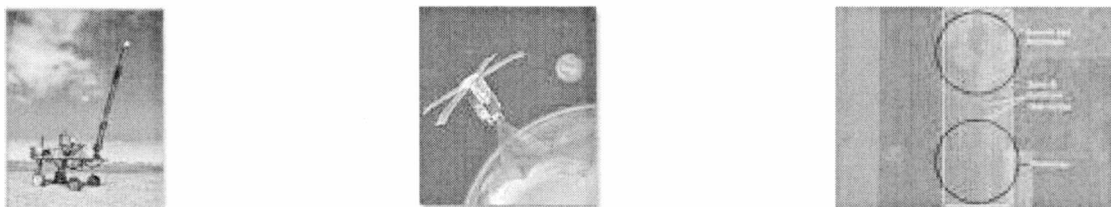


Imagen 4: Distintas técnicas de sistemas de sensores remotos

Los sistemas de sensores remotos permiten establecer el estado del suelo mediante el análisis de imágenes de la radiación emitida por éste y el cultivo. Estos sistemas pueden montarse en un satélite, un avión o hasta en un vehículo terrestre que recorra el campo. Ejemplos en Argentina de estas técnicas son las imágenes satelitales y las fotos aéreas (Bongiovanni y Lowenberg-DeBoer, 2001)

2.3.6. Sistemas de Información Geográfica (GIS)

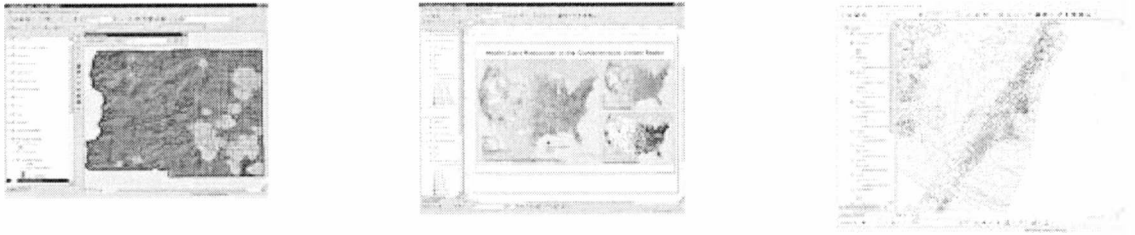


Imagen 5: Ejemplos de Sistemas de Información Geográfica

Los GIS se caracterizan por trabajar con datos que describen fenómenos que ocurren en la superficie terrestre (Gordillo 1998).

Un GIS es un sistema de gestión de bases de datos que permite almacenar, graficar, integrar y analizar información referenciada geográficamente (Baños y Goenaga, 2003). Longley (1999) describe estos sistemas desde tres puntos de vista; desde la visión de los mapas es un sistema para visualización y análisis de patrones, como base de datos administra eficientemente los datos geográficos para transacciones intensivas u operaciones complejas de análisis de datos y finalmente llega la visión del sistema cuyo objetivo es el análisis espacial y su definición es de sistema de información espacial.

La gran cantidad de datos georreferenciados que se recogen a lo largo de sucesivas campañas hace aconsejable utilizar los programas GIS (Baños y Goenaga, 2003). Estos son almacenados, combinados y analizados por estas herramientas facilitando su comprensión por parte de los productores permitiendo su utilización de manera productiva (Calp, 1999).

2.3.7. Tratamientos con Dosificación Variable (VRT)

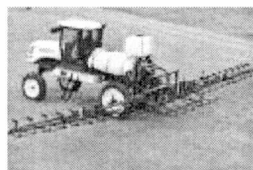
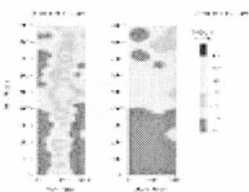


Imagen 6: Aplicación de insumos con dosificación variable

Son máquinas equipadas con hardware y software que controlan aplicaciones variables en tres áreas: siembra, pulverización y fertilización. Aplican la dosis específica de semilla, herbicida o fertilizante para cada punto del campo a partir de un mapa de prescripción generado por un GIS (Roberts, 2000).

Si no se dispone de un sistema de dosis variable automático, una alternativa es la dosis variable manual, o la “paralelización” de zonas de manejo.

2.4. Aspectos que deben tenerse en cuenta con la Agricultura de Precisión

La AP presenta algunas dudas a la hora de ser llevada a cabo por los productores, ya que en los lugares donde se utiliza en forma comercial no ha generado aumentos de rendimiento o reducción de insumos lo suficientemente importantes como para cubrir los costos extras. Esto se debe a que existe una considerable brecha de información entre los mapas de rendimiento y de muestreo de suelos y la aplicación de insumos. El problema con las recomendaciones de fertilización se debe a que los análisis económicos de beneficios y la estimación de la respuesta sitio-específica llevados a cabo no tienen en cuenta la estructura espacial de los datos (Bongiovanni, 2002b). Las decisiones de manejo óptimas dependen de la ubicación y características físicas del sector, del conocimiento que tiene el productor de esas características y de su capacidad para utilizarlo (Bullock et al., 2003).

Hay algunos problemas concernientes a la adaptación de la agricultura de precisión a las condiciones locales. Mientras el monitor de rendimiento en Argentina ha seguido el mismo camino de adopción que en Estados Unidos, la aplicación variable de insumos no ha sido ampliamente utilizada debido al alto costo del muestreo de suelos combinado con un relativo poco uso de fertilizantes. Una mayor adopción de tecnología de aplicación variable en las condiciones de Argentina dependerá del desarrollo de herramientas de manejo que usen información de bajo costo, tales como mapas de rendimiento, mapas de topografía, imágenes satelitales, fotografías aéreas y eventualmente sensores remotos y sensores de suelo (Bongiovanni, 2002a).

2.5. Análisis de Datos en la AP

El principal generador de muestras en la AP es el monitor de rendimiento. A pesar de la acumulación de datos de monitores de rendimientos por parte de productores, una queja común es que no existen buenas formas para analizar los datos. La comunidad de investigadores de la AP reconoce que trabajar con los datos de los monitores de rendimientos no es sencillo, pero si se demuestra el valor de los análisis entonces los desarrolladores de software encontrarán la manera de volverlo una tarea más fácil. (Erickson, 2005). Los investigadores y los productores pueden recolectar fácilmente enormes cantidades de datos pero transformar estos en decisiones de manejo significativas y evaluar riesgos y beneficios potenciales es una tarea que ha probado ser dificultosa (Dobermann et al., 2004).

2.6. Diferentes análisis de datos que se llevan a cabo en la AP

Se han adaptado técnicas de provenientes de la geología, economía regional, geografía y otras disciplinas con mayor pasado en el análisis de datos espaciales y se han aplicado a problemas específicos de los datos de la AP (Erickson, 2005).

Los tipos de análisis pueden ser baja complejidad como identificar zonas “a ojo” en los mapas de rendimientos y de topografía para realizar tratamientos diferenciados en base a recomendaciones agronómicas para la zona en análisis o de alta complejidad y con rigor científico con un diseño experimental (Bongiovanni, 2002a; Lowenberg-DeBoer and Swinton, 1997; Anselin et al., 2004; y otros.)

Como ejemplo y guía de análisis, este trabajo toma la tesis doctoral de Bongiovanni (2002a). El trabajo se basa en aplicar los métodos de la econometría espacial para estimar una función de respuesta en un modelo de rendimiento. La econometría espacial está orientada hacia la estimación y la comprobación de hipótesis. Asume que la covarianza espacial es una interacción entre objetos espaciales discretos y requiere la especificación de un procesos estocástico espacial con una estructura de observaciones contiguas (matriz de ponderaciones espaciales).

En este trabajo el análisis se aplica a datos espaciales con una resolución similar al ancho del cabezal de la cosechadora.

La econometría espacial tiene en cuenta los efectos de la auto correlación espacial en forma explícita, permite la utilización de técnicas de bajo costo para corregir la heterocedasticidad y así estima coeficientes más eficientes, teniendo en cuenta el grado en que las diferentes regiones topográficas dentro del lote afectan la magnitud, la significancia y el signo de los coeficientes estimados.

La idea es medir el impacto de la mayor precisión que se obtiene al usar la econometría espacial en un análisis económico basado en presupuestos parciales.

Finalmente se concluye que econometría espacial se puede aplicar al estudio de los datos del monitor de rendimiento para estimar la respuesta del cultivo al fertilizante, en este caso la respuesta del maíz al N. La metodología es flexible y tiene un gran potencial para el diseño de ensayos a campo.

2.7. Metodologías y herramientas utilizadas en el análisis de la AP

Software para mapeos básicos pueden ser incluidos de manera gratuita o por un pequeño monto con la compra de un monitor de rendimiento. Los sistemas de información geográficos que permiten análisis de datos y proveen herramientas para tomar decisiones pueden requerir una inversión financiera considerable. Este tipo de servicios se encuentra disponible a partir de proveedores de servicios para la agricultura (Casady et al., 1998).

A lo largo de un proceso de análisis y dependiendo de las características del mismo normalmente se vuelve necesario la utilización de distintas herramientas de software como planillas de cálculo, paquetes estadísticos y sistemas de información geográfica para modelar o convertir datos. Esto puede se puede observar en el trabajo de Bongiovanni (2002a) y en las recomendaciones de Griffin et al. (2005), entre otros.

2.8. Soluciones de Software existentes

Los paquetes existentes trabajan en su mayoría sobre GIS comerciales que están contruidos en modo completamente propietario obligando a tener que adquirir el

software entero con funciones que nunca se usaran y significando un alto costo, manteniendo características de interfaz que no hacen al dominio y dificultan la claridad, no favorecen la integración con otros sistemas, no brindan información completa sobre los métodos que utilizan para la manipulación de los datos ni permiten su adaptación (Saraiva 1998).

El usuario termina utilizando múltiples productos como GIS, planillas de cálculo, programas de cálculo estadístico, etc., realizando conversiones y manipulación de datos que incomodan la actividad.

2.9. Problemas de dominio de GIS

La tecnología de GIS involucra tratar con aspectos complejos que van desde la adquisición de datos, la corrección de los mismos, la representación de relaciones espaciales y de la topología y el diseño de la interfaz. Uno de los problemas más graves que existen en este dominio es la carencia de un método de diseño que permita reflejar todas las características mencionadas y permitir obtener un producto con las propiedades deseables desde el punto de vista de la ingeniería de software, como por ejemplo, reusabilidad, modularidad, facilidades para la evolución y la modificación, etc.

En efecto, hoy en día la construcción de este tipo de aplicaciones se vuelve una tarea completamente artesanal en donde aquellos diseñadores experimentados utilizan su conocimiento anterior “on the fly” mientras que aquellos que no poseen experiencia realizan la tarea basándose en las “pruebas y errores”. Por supuesto, es sumamente difícil encontrar documentación sobre estos sistemas, pues la preocupación pasa más por la precisión y la corrección de los datos que se están almacenando y en resolver las dificultades que presentan las relaciones espaciales, que por obtener un buen diseño.

Esta metodología lleva a lo que ya se ha conocido y vivido en otras áreas, los costos de mantenimiento, reusabilidad y evolución de los sistemas resultan altísimos. De hecho, ocurre más de una vez, que grupos que realizan desarrollos basados en el mismo tipo de aplicación (por ejemplo catastral) deben comenzar cada trabajo como si fuera completamente nuevo, cuando en realidad la mayor parte de los módulos de una aplicación catastral pueden ser utilizados para otra. (Gordillo 1998).

2.10. Tecnología Orientada a Objetos (OO)

La Orientación a Objetos es la organización de la solución de software a un problema como una colección de objetos discretos que incorporan tanto estructuras de datos como comportamiento. Los principales aspectos que involucra esta aproximación son la clasificación, la herencia y el polimorfismo.

- **Clasificación:** Los objetos con los mismos atributos y el mismo comportamiento se agrupan dentro de una clase. Cada clase describe un conjunto posiblemente infinito de objetos individuales. Los objetos son llamados instancias de una clase.
- **Herencia:** Es la capacidad de compartir atributos y comportamiento entre clases basada en una relación jerárquica. Una clase ya definida puede ser refinada en subclases, cada subclase incorpora o hereda todos los atributos y el comportamiento de su superclase y agrega los suyos propios.
- **Polimorfismo:** Significa que la misma operación puede comportarse de manera diferente en distintas clases. Una implementación específica de una operación en una cierta clase es llamado método. Al ser OO polimórfico puede existir más de un método que implemente una operación (Rumbaugh et al., 1991).

2.11. Beneficios de la Tecnología Orientada a Objetos

La tecnología orientada a objetos ha demostrado ser una buena herramienta a la hora de resolver problemas de gran envergadura, en los que la complejidad de la información que se manipula es alta y es imprescindible contar con un alto grado de integración de esa información. Como resultado de su uso se obtienen, no solamente soluciones en donde las características anteriores se mantienen (reusabilidad, evolución, modularidad) sino también sistemas interoperables, ya que los objetos pueden encapsular conocimiento que puede agruparse según las necesidades (Gordillo, 1998).

2.12. Problemas con las soluciones OO

La preocupación de los diseñadores de aplicaciones geográficas por obtener información correcta desde el punto de vista de la representación (precisión, escala, etc.) genera un error bastante común en el diseño que parte de considerar las características de la representación espacial de las entidades, en lugar de pensarlas en función de su definición abstracta. El énfasis entonces está puesto en cómo la información se representa y no en qué información se debe representar. Como una consecuencia de esto, las entidades que tienen distintas representaciones (por ejemplo distintas escalas) se implementan usualmente como objetos diferentes cuando en realidad no lo son (Gordillo, 1998).

Existen muchos trabajos que se han abocado al estudio del uso del modelo de objetos como una solución a estos problemas. Sin embargo, la mayoría de estos trabajos se basan en el modelo para representar entidades complejas pero no han estudiado cuestiones tales como la separación de las representaciones físicas de los objetos (puntos, líneas, polígonos) de su comportamiento conceptual, lo que fuerza al diseñador a asignar prematuramente características espaciales a las entidades y hace difícil trabajar con diferentes representaciones del mismo objeto (Gordillo, 1998). Como ejemplo se puede citar el trabajo de Saraiva et al. (1998).

2.13. Unified Modeling Language (UML)

El estándar UML ayuda a especificar, visualizar, y documentar los modelos de sistemas de software, incluyendo su estructura y diseño, de manera tal de reunir todos estos requerimientos (Siegel, 2005).

2.14. Patrones de Diseño OO

Los patrones de diseño describen soluciones simples y elegantes a problemas específicos del diseño OO. Capturan soluciones que han sido desarrolladas y evolucionadas a lo largo del tiempo. Por lo tanto no son diseños que la gente tiende a generar cuando se enfrenta inicialmente a un problema. Estos reflejan rediseños y remodelación que diseñadores

han generado buscando una mayor reusabilidad y flexibilidad en su software. Los patrones de diseño capturan estas soluciones en una forma sucinta y fácilmente aplicable. Los patrones de diseño facilitan la reusabilidad de diseños y arquitecturas exitosas. El hecho de expresar técnicas ya probadas como patrones de diseño las vuelve más accesibles a desarrolladores de nuevos sistemas. Los patrones de diseño ayudan en la elección de alternativas para hacer un sistema reusable y evitan las alternativas que no favorecen esta reutilización (Gamma et al., 1995).

Capítulo 3: Etapas y características de un proceso de análisis de datos de un monitor de rendimiento.

Un análisis en particular puede comprenderse como una serie de procesos al que se somete cierto conjunto de muestras. Varios de estos procesos son comunes a los distintos métodos de análisis que se llevan a cabo en la actualidad como el filtrado de datos, la determinación de zonas de manejo y otros.

A continuación se describen los distintos procesos que se han identificado en la bibliografía abarcada y se propone una conclusión para cada uno de ellos que detalle los elementos relevantes para el proceso de definición del modelo.

3.1. Filtrado de Datos

Es un proceso de preparación del conjunto de muestras del monitor de rendimiento llamado “limpieza de datos” o “filtrado de datos” aunque en realidad no es mucho más que remover mediciones que son identificadas como erróneas debido a la dinámica de la cosechadora y al comportamiento del operador de la misma. Este proceso de filtrado de datos no significa de ninguna manera una modificación o manipulación de los datos. Sin embargo ésta incrementa la calidad del conjunto de datos (Griffin et al., 2005). Por ejemplo se puede llevar a cabo una eliminación de los datos de los extremos del lote, como así también los excesivamente altos o bajos usando un criterio estadístico (Bongiovanni, 2002a).

De lo antes expuesto se desprende que el filtrado de datos es un proceso en el que se descartan muestras a partir de algún criterio definido, como la pertenencia a los extremos del lote donde la variación de la velocidad de la cosechadora y las maniobras necesarias generan lecturas erróneas, criterios estadísticos como búsquedas de outliers para determinar fallas del muestreo, etc. Podrían ser múltiples y combinarse para lograr el resultado deseado.

3.2. Modelado de Datos

Raramente las diferentes fuentes de datos comparten la misma densidad (resolución espacial), por lo tanto es necesario algún tipo de modelado de datos para unificar la densidad. Hay que ser cuidadosos con el proceso que se elige para minimizar la interferencia. Los datos del monitor de rendimiento son usualmente los más densos, los datos provenientes del muestreo de suelos son los menos densos. La práctica habitual que recomienda Anselin (2001) citado por Griffin et al. (2005) es la de mantener el conjunto de datos menos denso como la base para el resto de las capas de datos cuidándose de no llevar a cabo métodos de interpolación vía kriging o métodos geoestadísticos para no generar interferencia.

Una opción es una modelización en polígonos cuadrados del ancho del cabezal de la cosechadora, y asimilando el resto de los conjuntos de datos a partir de sus respectivos promedios dentro del polígono (Bongiovanni, 2002a). Debido a que los datos del monitor de rendimiento incluyen observaciones más cercanas que la distancia entre hileras, estos datos se promedian creando cuadrículas para que la distancia dentro de la hilera sea equivalente a la distancia entre hileras.

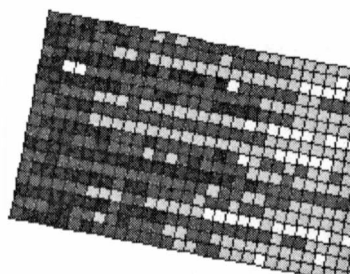


Imagen 1: Modelado en cuadrículas en las que se promedia el rendimiento del cultivo

Se asume, a partir de la bibliografía, que el modelado de datos en estructuras espacialmente uniformes permite llevar a cabo los análisis estadísticos y económicos con una base de datos balanceada. La superficie del lote se divide en elementos discretos donde se agregan los datos permitiendo integrar datos de diferentes orígenes y resoluciones espaciales. Como ejemplo se pueden generar cuadrículas en las que se promedian los valores de las muestras que caen dentro de cada una de ellas.

3.3. Definición de Zonas de Manejo

En el manejo de la variabilidad se utiliza habitualmente el concepto de “zonas de manejo” (ZM), que se definen como sectores dentro de los lotes que expresan una combinación relativamente homogénea del rendimiento posible del cultivo, eficiencia de uso de los insumos y del impacto ambiental (Gregoret et al., 2005).

Para identificar las zonas de manejo se utilizan herramientas informáticas como el Management Zone Analyst, (Mizzou-ARS, 2000). Las áreas de manejo se definen mediante información previa para agrupar datos espaciales que poseen información multidimensional. Este programa hace una clasificación difusa no supervisada (“*unsupervised fuzzy classification*”) para crear zonas de manejo potenciales basándose en la información disponible. Se determinan zonas dentro del ensayo, en forma previa a todos los análisis de datos.

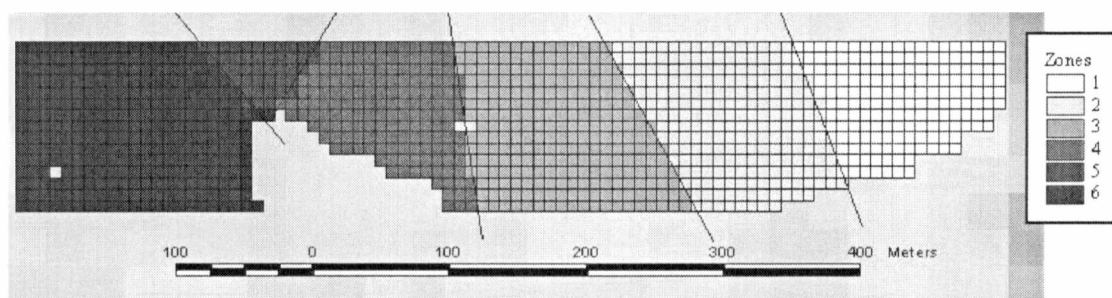


Imagen 2: Determinación de zonas de manejo

Se deduce en las definiciones anteriores que la intención en la determinación de zonas de manejo es generar particiones del lote donde exista cierta uniformidad de manera de simplificar el análisis al posibilitar la generación de estimaciones para cada una de estas zonas. Es necesario indicar la cantidad de zonas en las que se desea dividir las muestras y deben ser comparables entre sí las distintas instancias para poder determinar el número con mejor sentido.

3.4. Agregación de información sobre tratamientos

La información sobre tratamientos se agrega al conjunto de datos, estos pueden ser tratamientos de laboreo, diseños de ensayos a franjas, etc. Se pueden crear distintos

polígonos y éstos ser compuestos para formar un mapa de tratamiento. Esto mismo se puede hacer con variables dummy para tipos de suelo, etc. (Griffin et al., 2005). Esta información extra debe asimilarse al modelo de datos elegido para llevar adelante el análisis.

Por ejemplo, Bongiovanni (2002a) utilizó un diseño experimental en franjas de dosificación de Nitrógeno, y estos datos fueron luego agregados a la capa formada por los polígonos.

3.5. Cálculos adicionales con planillas de cálculo

Las planillas de cálculo tipo MS Excel resultan útiles para calcular variables adicionales, estas variables pueden incluir variables dummy, identificadores únicos si aún no fueron creados, etc. (Griffin et al., 2005). Algunos GIS cuentan con una capacidad de análisis similar a las planillas de cálculo.

3.6. Análisis estadísticos

Existen diversas formas de llevar a cabo un análisis estadístico de los datos de un monitor de rendimiento con diferencias en la eficiencia de los estimadores alcanzados.

3.6.1. Problemas con análisis estadísticos tradicionales:

El problema con la estadística tradicional es que asume que las observaciones son independientes y no tiene en cuenta los patrones espaciales que pueden ser encontrados en los datos, estos patrones son importantes en los residuos de la regresión y al ignorarlos vuelven ineficientes los estimadores y el error estándar se ve influenciado (Bongiovanni 2002a, Capítulo 2).

Los análisis estadísticos tradicionales como ANOVA y el método de regresión de mínimos cuadrados ordinarios (OLS) no son confiables en presencia de variabilidad espacial. Los análisis de regresión espacial son una metodología que se sobrepone a estas limitaciones de los análisis tradicionales (Griffin et al., 2005).

3.6.2. Técnicas de regresión espacial en el análisis estadístico:

Análisis estadísticos espaciales con diferentes modelos de correlación entre observaciones vecinas han sido desarrollados en diferentes contextos (Geografía,

Economías Regionales, Geología, etc.). Lambert et al. (2004) enumeran cuatro técnicas de regresión espacial y las aplican al análisis de los datos del monitor de rendimiento, realizando una comparación entre ellas y OLS también; las mismas son econometría espacial, regresión polinomial con tendencia, análisis de vecinos cercanos y geoestadística. Concluyendo que la econometría espacial es el método con mayor significancia estadística.

3.6.3. Econometría espacial:

La econometría espacial asume que la variabilidad espacial es una relación entre observaciones discretas. La estructura espacial puede ser fundada tanto en la variable dependiente o en el error residual. Esta estructura espacial es modelada asumiendo que la variable dependiente o el error residual están en función de la suma de los datos de sus vecinos (Anselin, 1988).

Estas dos formas de explicar la relación espacial pueden conducir a usar uno u otro modelo de análisis:

- “*de ponderación espacial*” (*variable dependiente correlacionada espacialmente*)

$$y = \rho W y + X \beta + \varepsilon$$

donde y es un vector (n por 1) de observaciones de la variable dependiente, ρ es el coeficiente auto regresivo, $W y$ es el vector de variables independientes ponderado por una matriz W de observaciones vecinas, X la matriz de observaciones de las variables independientes (n por K) y ε es el error de muestreo.

- “*de error espacial*” (*error de muestreo correlacionado espacialmente*)

$$y = X \beta + \varepsilon \quad \text{donde } \varepsilon = \lambda W \varepsilon + u$$

donde u es el error de muestreo aleatorio que sigue una especificación auto regresiva espacial (SAR) con un coeficiente auto regresivo λ .

Ambos tipos de auto correlación espacial se pueden corregir estimando los coeficientes de correlación espacial, lo que requiere el uso de una matriz de ponderaciones espaciales W , que pondera los datos de acuerdo a la contigüidad entre observaciones (Bongiovanni, 2002a).

El modelo de regresión elegido por Bongiovanni (2002a) es el “**de error espacial**”. Estima la respuesta del maíz al nitrógeno (N) como una función cuadrática en N por posición topográfica dentro del lote, con i señalando la posición topográfica (i = Bajo, Pendiente, y Loma) y j señalando la ubicación geo-referenciada dentro de cada posición topográfica:

$$Rendimiento_{ij} = \alpha_i + \beta_i N_{ij} + \gamma_i N_{ij}^2 + \delta \text{ posición} + \varepsilon_{ij}$$

Donde $Rendimiento_{ij}$ es el rendimiento del maíz (de un monitor con GPS), N_{ij} es la dosis de N aplicada en cada franja del ensayo, posición se refiere a las variables dummy por posición topográfica, y ε_{ij} son los términos de error de la regresión por la ubicación. Esta especificación permite la estimación del efecto de las diferentes posiciones topográficas dentro del lote, α_i (como diferencia del efecto medio) como así también de los términos de interacción entre posición (δ), N (β_i) y N^2 (γ_i).

El objetivo primordial del análisis de regresión dentro de la econometría espacial es estimar el valor de una variable aleatoria (la variable dependiente: el rendimiento), dado que los valores de una o más variables asociadas (la o las variables independientes: dosis de N, tipo de suelo, elevación, etc.) son conocidos. La ecuación de regresión es la fórmula algebraica por la cual se determina el valor estimado de la variable dependiente o de respuesta.

La principal contribución de este estudio es el demostrar que la econometría espacial se puede aplicar al estudio de los datos del monitor de rendimiento para estimar la respuesta del cultivo al fertilizante, en este caso la respuesta del maíz al N. La metodología es flexible y tiene un gran potencial para el diseño de ensayos a campo.

3.7. Interpretación de los resultados

Los estimadores que se han obtenido son utilizados con fines económicos para determinar dosis óptimas de aplicación, su rentabilidad y la generación de recomendaciones de dosis sitio-específicas.

Bongiovanni (2002a) determina la rentabilidad de la utilización de dosis variable de N con respecto a lo que un prestador de servicios cobraría “extra” por aplicar con dosis variable.

El análisis económico se realiza usando el método de presupuestos parciales que determina si los beneficios superan los costos variables en la campaña del cultivo. El retorno neto al N se calcula usando análisis marginal, en el que la dosis óptima de N que maximiza la rentabilidad se identifica como el punto en el que el valor de rendimiento extra que se logra es igual a esa unidad de N (MVP = MFC). Si se usa más que eso, se pierde dinero en cada unidad de N que se aplica, porque el incremento de rendimiento no paga el costo del N aplicado. Estas dosis óptimas “económicas” se comparan con la dosis “agronómica” sugerida para la zona de Río Cuarto.

El problema de maximización de la rentabilidad esperada para un año se expresa, por ejemplo, como:

$$\text{Max } E[\pi] = \sum_{i=1}^4 \text{Area}_i * E \left[P_c * (\alpha_i + \beta_i * N_i + \gamma_i * N_i^2) - r_N * N_i \right]$$

donde:

E = Esperado

π = Retornos netos totales por sobre el fertilizante N (\$ ha⁻¹)

Area_i = Proporción de área j ($j = 1, \dots, m$)

i = Área: 1=Bajo Este, 2=Pendiente Este, 3=Loma, 4=Pendiente Oeste

P_c = Precio del maíz a valores año 2001 (\$8,20 por quintal o por 100Kg)

α_i = Coordenada al origen del modelo auto regresivo espacial

β_i = Coeficiente estimado lineal

γ_i = Coeficiente estimado cuadrático

N_i = Cantidad de N elemental aplicado en área i

r_N = Precio del N elemental a valores año 2001 ($\$0,61 \text{ Kg}^{-1}$), más intereses por seis meses a una tasa del 15% anual (Precio de la urea: $\$0,26 \text{ Kg}^{-1}$)

Los análisis económicos brindan indicadores útiles para el análisis económico de la explotación, estos indicadores obtenidos se utilizan para justificar la implementación de tratamientos de dosificación variable, estimar sus beneficios y sus costos, etc.

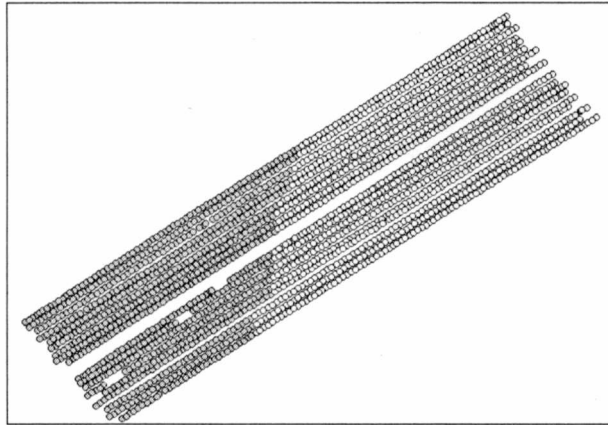


Imagen 3: Ejemplo de recomendaciones de aplicación variable para dos zonas de manejo determinadas

También se generan mapas de aplicación para ser utilizados por las herramientas de dosificación variable. El resultado son mapas de aplicación de insumos que se cargaran en las herramientas durante la próxima campaña buscando maximizar el rendimiento.

Capítulo 4: Arquitectura de la estrategia de análisis de datos

En este capítulo se propone una arquitectura abstracta que resuelva el análisis de datos en la AP contemplando las distintas opciones disponibles en cada uno de los procesos de análisis que se aplican a los datos de entrada.

4.1. Arquitectura del sistema

Las actividades que se llevan a cabo en el dominio se pueden dividir en tres etapas, primero la recolección de muestras, luego el modelado y análisis de datos y finalmente la utilización de los resultados. Esto sugiere la estructura general del dominio

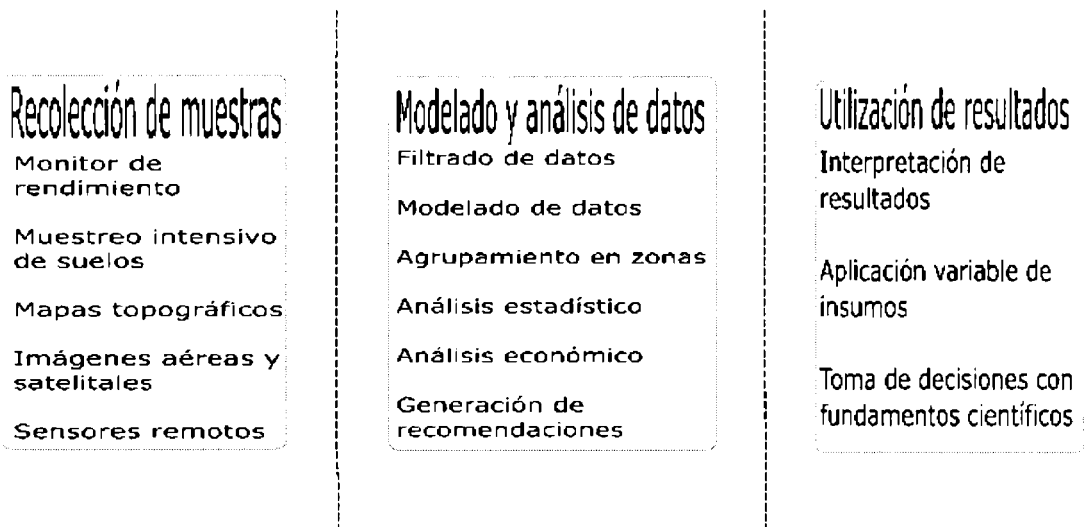


Imagen 1: Arquitectura del dominio

Siendo la etapa de modelado y análisis de datos la de interés en este trabajo y sobre la que se desarrollará el análisis exhaustivo para generar la arquitectura que lo resuelva.

4.2. Modelado y análisis de datos

El análisis de datos en la AP se puede definir como una serie de procesos que se llevan a cabo sobre el conjunto de muestras que se disponen de una o más campañas acerca de un lote específico.

El origen de las muestras son principalmente los datos recogidos por el monitor de rendimiento montado en las cosechadoras y datos provenientes de muestreo intensivo de suelos. También es posible trabajar con mapas topográficos que describen el paisaje, fotos aéreas e información relevada por sensores remotos sobre la composición del suelo.

Existen diversos análisis que se utilizan en la actualidad y seguramente se diseñarán nuevos en el futuro, en general los análisis que se han abarcado en esta investigación comparten procesos genéricos que se aplican a los datos; como filtrado de muestras, modelado en elementos discretos, estimación de variables por medio de análisis estadísticos, análisis económicos para determinar dosis óptimas, y generación de recomendaciones, pero se diferencian en las actividades específicas que se realizan en cada una de ellos.

Cada uno de estos análisis significa una estrategia distinta en la forma de la selección, configuración y composición de los procesos mencionados, incluyendo que en un diseño de análisis específico es posible el intercambio de procesos y configuraciones de estos para obtener variantes de los análisis comparables entre sí mediante criterios definidos por el usuario.

4.2.1. El Lote

Los análisis son llevados a cabo sobre un lote específico dentro de un establecimiento productivo. El lote es una porción de superficie terrestre con límites geográficos establecidos. Las características del paisaje del mismo se obtienen mediante el análisis de información topográfica determinando zonas como bajo, loma o pendiente. Asimismo, según Hatfield (2000), la variabilidad de los rendimientos dentro de un lote de producción posee tres componentes: i) natural (tipo de suelo y topografía); ii) aleatoria (precipitaciones); y iii) de manejo (aplicación de fertilizantes, densidad de siembra, etc.). La interacción entre estas tres fuentes de variabilidad genera efectos que no siempre están en concordancia con los límites establecidos para un lote de producción. Una de las finalidades de la AP es identificar diferentes ambientes dentro de los lotes que poseen cualidades homogéneas determinadas principalmente por topografía y condiciones del suelo, y que se denominan “zonas de manejo” (ZM).

4.2.2. Las Campañas

Las actividades que se llevan a cabo en un Lote se agrupan en Campañas (años), para cada una de estas campañas se dispone información sobre tratamientos como aplicación de fertilizantes, de herbicidas, muestreo intensivo de suelos, datos del monitor de rendimiento, etc.

4.2.3. Estrategias de Análisis:

Las estrategias de análisis se componen de una serie de procesos que se aplican a muestras recogidas por distintos medios con el fin de explicar el comportamiento del cultivo a la aplicación de algún insumo en particular. Por ejemplo se puede intentar explicar la respuesta a la aplicación de Nitrógeno en un cultivo de Maíz a partir de las muestras recogidas por un monitor de rendimiento, información del paisaje y con información de un diseño experimental en franjas de la aplicación de Nitrógeno en la campaña estudiada con el fin de determinar la dosis económicamente óptima y analizar su rentabilidad. A partir de las conclusiones es posible generar un mapa de recomendaciones para aplicaciones de insumos en futuras campañas.

4.2.4. Los Procesos

Los procesos son módulos de transformación de datos que publican una interfaz específica definiendo su entrada/salida.

Cada uno de estos procesos se lleva a cabo con determinada configuración y debe ser posible almacenar sus resultados.

Es importante definir métricas para comparar las posibles variantes a los distintos procesos.

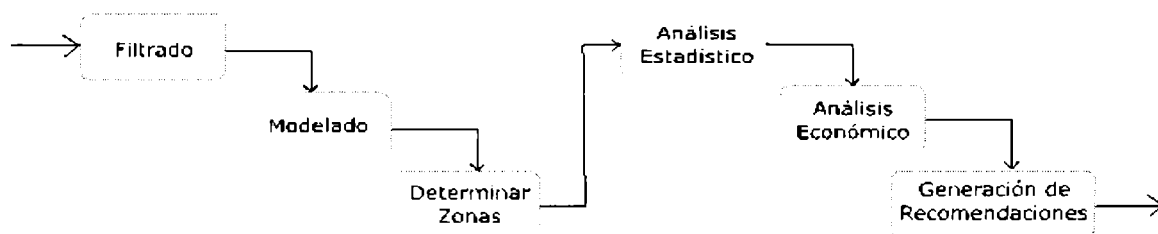


Imagen 2: Composición de Procesos

4.3. Análisis de Casos de Uso

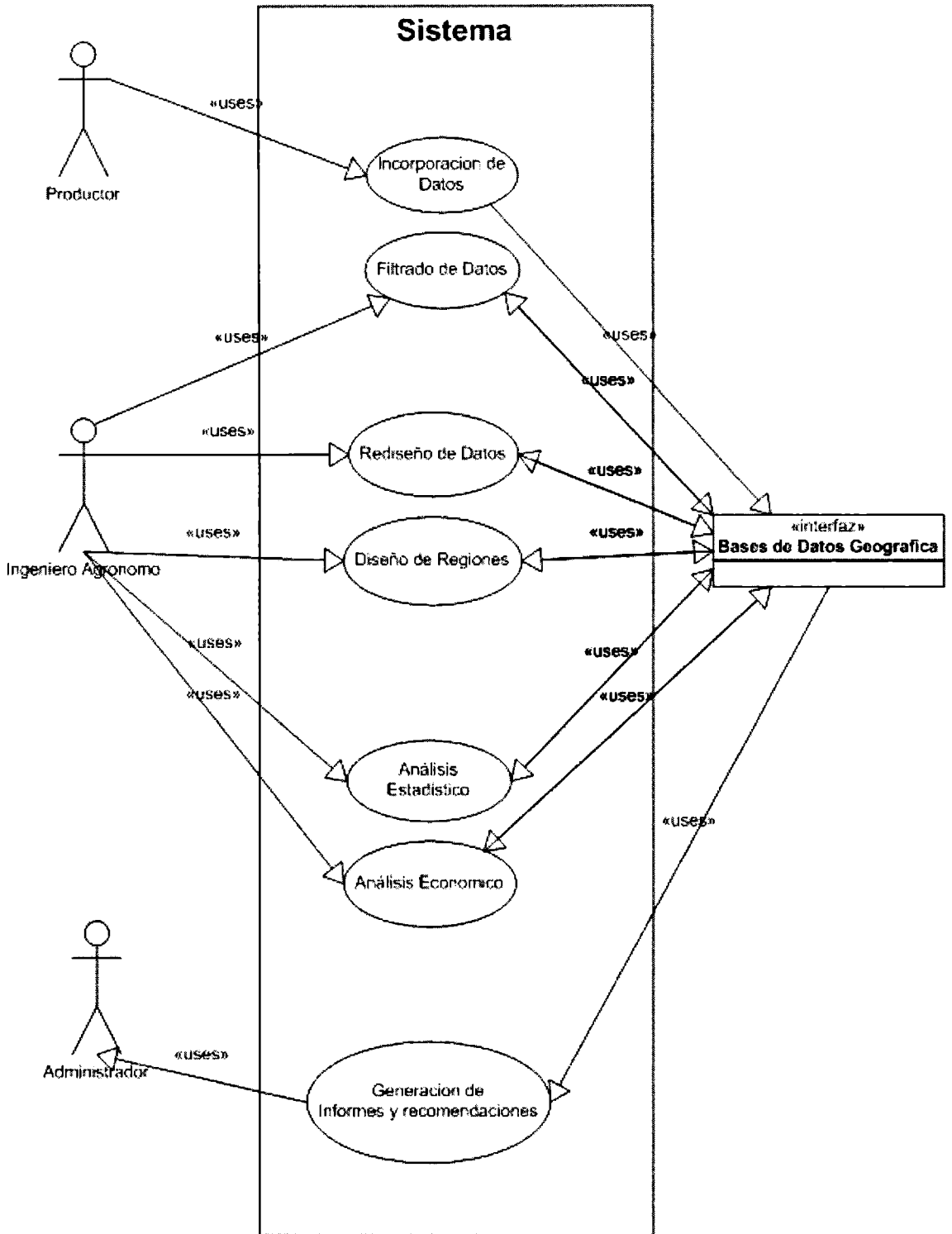


Imagen 3: Casos de uso

4.4. Metodología

Se trabajará analizando cada elemento que se ha identificado en el dominio utilizando para la generación de resultados modelos UML. En casos donde se identifique alguna analogía con un Patrón de Diseño OO se introducirá una instancia de este justificando la decisión tomada llegando finalmente mediante este proceso al modelo final.

4.5. Resultado

En este capítulo llevaremos a cabo la metodología enumerada en el capítulo anterior, se analizará cada elemento y se propondrá el modelo correspondiente.

4.5.1. Modelo para incorporación de características geográficas a elementos del dominio

Para incorporar características geográficas a los elementos del dominio emplearemos el modelo propuesto por Gordillo (1998) que desacopla la representación geográfica del elemento.

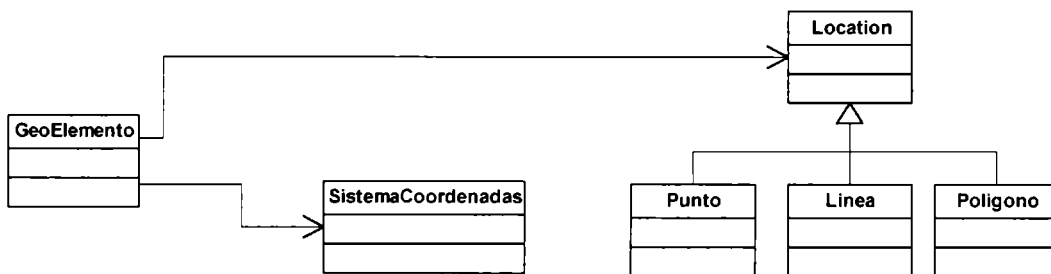


Imagen 4: Modelo utilizado para incorporar las características geográficas de los elementos

Esto lo logra al independizar al elemento de dominio de su representación como Punto, Línea o Polígono y abstrayendo también el Sistema de Coordenadas con el que se representa la ubicación de los elementos en superficie terrestre.

4.5.2. El Lote y las Campañas

El Lote posee un nombre natural determinado por el explotador del terreno, los límites los modela con una agregación con la clase GeoElemento que independiza la representación geográfica del elemento de dominio y conoce la información de las campañas con una asociación de cardinal uno a muchos con la clase Campaña, gracias a lo cual mantiene el historial de trabajos y datos muestreados.

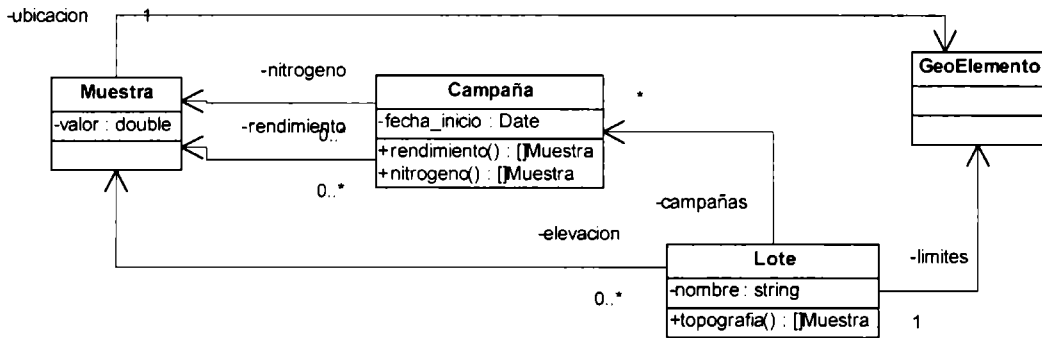


Imagen 5: Un Lote y sus Campañas

4.5.3. Estrategias de Análisis:

patrón Strategy

Propósito:

Define una familia de algoritmos, encapsulando cada uno de estos, haciéndolos intercambiables. Esto permite cambiar el algoritmo independientemente de los clientes que lo usan.

Estructura:

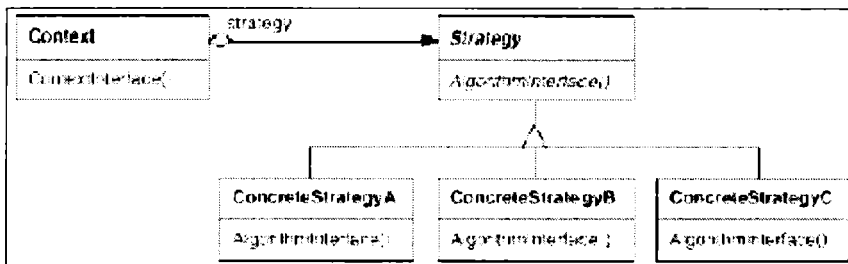


Imagen 6: Estructura del patrón Strategy

Participantes:

- *Strategy*: declara una interfaz común a todos los algoritmos soportados. El *Context* usa esta interfaz para invocar al algoritmo definido por alguna *ConcreteStrategy*.
- *ConcreteStrategy*: implementa el algoritmo utilizando la interfaz de *Strategy*.
- *Context*: es configurado con algún objeto *ConcreteStrategy*, mantiene referencia a este objeto, podría definir una interfaz para que el *Strategy* acceda a sus datos

Instancia del patrón Strategy modelando la Estrategia de Análisis:

El modelar la Estrategia de Análisis utilizando el patrón Strategy nos permite intercambiar los distintos análisis existentes sin que impacte esto sobre la representación de los Lotes, permitiendo simultáneamente extender con poco esfuerzo la familia de análisis disponibles aprovechando los beneficios de la herencia y el polimorfismo que se han enumerado en el capítulo 3.10 y que favorecen el intercambio de objetos de manera transparente al resto de los elementos involucrados.

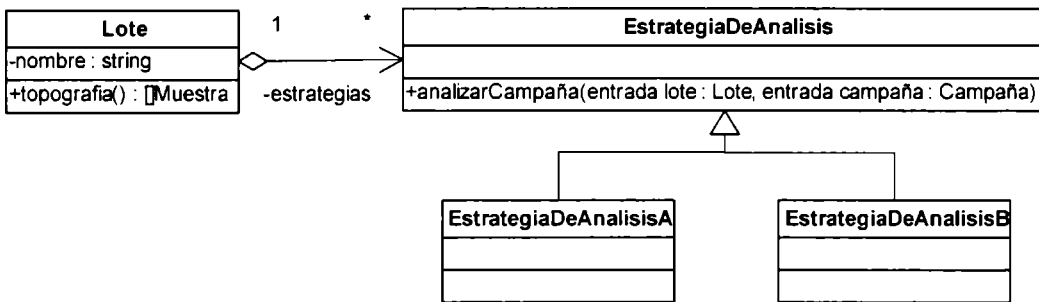


Imagen 7: Instancia del patrón Strategy para el Análisis

4.5.4. Los Procesos

Los procesos en general serán instancias del patrón Strategy, logrando de esta manera ser intercambiables independientemente de la estrategia que los utiliza, favoreciendo la incorporación de nuevos y distintos procesos en el futuro.

4.5.4.1. Filtrado de Datos

La interfaz definida para el Filtro permite su composición en secuencias de manera de lograr el resultado final esperado. Un filtro puede recibir como entrada de datos la salida de un filtro anteriormente aplicado.

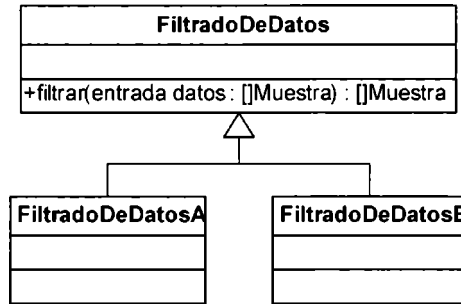


Imagen 8: Filtrado de datos

4.5.4.2. Modelado

En el Modelado se posibilita la incorporación de múltiples fuentes de muestras y datos que serán incorporados a cada uno de los elementos discretos resultantes por medio del criterio establecido.

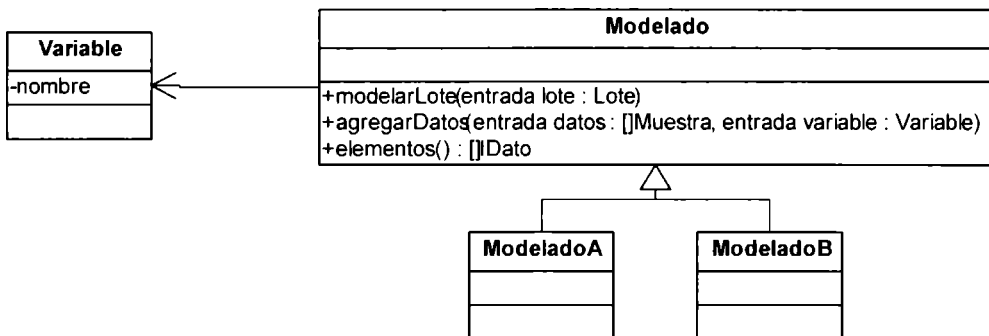


Imagen 9: Modelado

Recordemos que es necesaria una transformación de la naturaleza continua del dominio en elementos discretos, de esta manera en esta etapa se incorporan en forma de promedios o de alguna otra las muestras que fueron recolectadas con distinta densidad como lo describimos en el capítulo 3.2.

4.5.4.3. Determinación de Zonas de Manejo

El proceso de Determinación de Zonas posibilita la elección del número de zonas que serán creadas para luego comparar entre los resultados obtenidos las zonas que mejor representan el lote en función a los elementos que las componen. Como resultado se obtienen Zonas de manejo que son identificadas por un nombre, y una representación geográfica independiente del modelo de dominio.

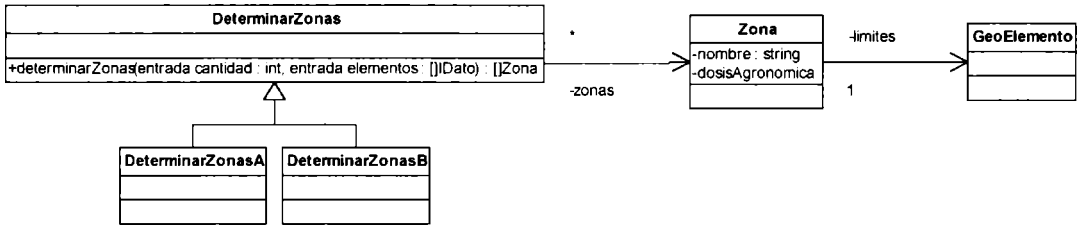


Imagen 10: Determinación de zonas de manejo

La determinación de zonas de manejo agrupa los elementos discretos previamente creados en zonas uniformes mediante algún criterio, de esta forma queda la superficie original del lote dividida en n particiones determinadas por el usuario.

4.5.4.4. Análisis Estadístico

En esta etapa se generan estimaciones sobre cada una de las variables involucradas en el análisis con sus correspondientes indicadores de eficiencia como p-valor. Los mismos serán utilizados en el futuro análisis económico.

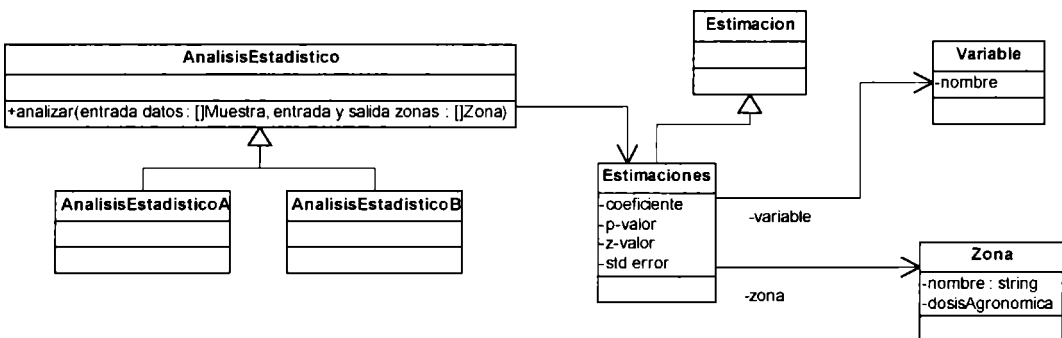


Imagen 11: Análisis estadístico

Los Análisis Estadísticos explotan las características de ser instancia del patrón Strategy logrando reducir la complejidad de incorporar nuevos análisis al espectro disponible y favoreciendo la comparación entre ellos por medio de indicadores de confianza como el p-valor de los coeficientes estimados.

4.5.4.5. Análisis Económico

En el análisis económico se sacan conclusiones sobre los beneficios económicos de aplicar una u otra estrategia, o simplemente el costo de no aplicar ninguna estrategia de dosificación variable. Como estimadores se trabaja con la dosis óptima agronómica y la dosis óptima económica obtenidos a partir de las estimaciones del análisis estadístico sobre el cual se llevara a cabo el análisis económico en particular.

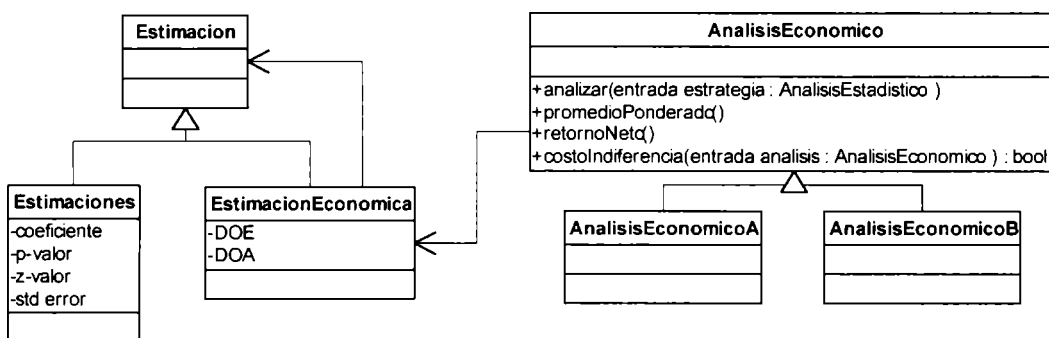


Imagen 12: Análisis económico

Aquí también se aprovechan las características del patrón Strategy explotando la capacidad de realizar múltiples análisis económicos para fundamentar por todos los medios disponibles las decisiones que serán tomadas en futuras campañas.

Las Estimaciones Económicas agregan información a una estimación económica, tienen la forma del patrón de diseño Decorator.

Patron Decorator

Propósito:

Agrega responsabilidades adicionales a objetos dinámicamente. El patrón Decorator provee una alternativa flexible a la subclasificación a la hora de extender funcionalidad.

Estructura:

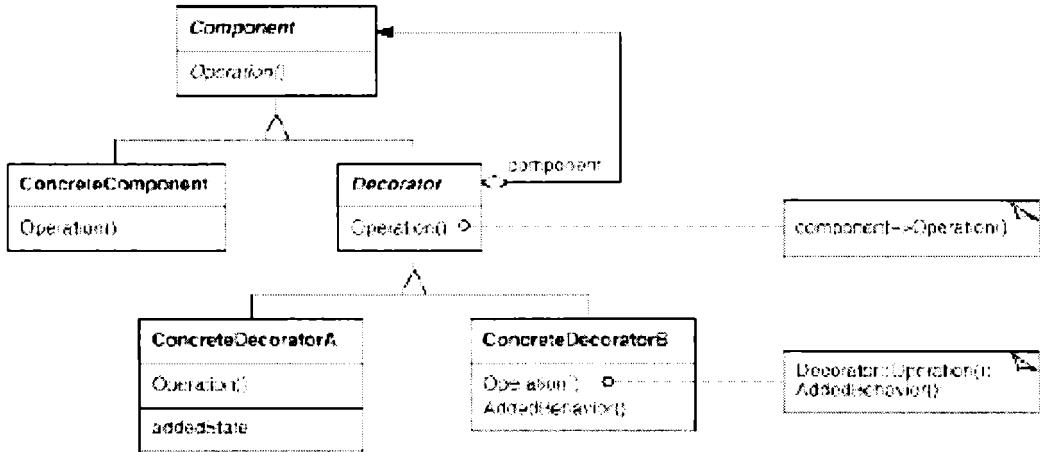


Imagen 13: Estructura del patrón Strategy

Participantes:

- *Component*: Define la interfaz de los objetos que poseen responsabilidades adicionales dinámicamente.
- *ConcreteComponent*: Define un objeto al que se le puede adicionar responsabilidades dinámicamente.
- *Decorator*: Mantiene la referencia a un objeto Component y define una interfase conforme a la interfase Component
- *ConcreteDecorator*: Adiciona responsabilidades a un Component

4.5.4.6. Generación de Recomendaciones

Las recomendaciones son mapas de aplicación con los que se programaran las herramientas con dosificación variable y que les indicaran la cantidad de insumo especifica que se aplicara a cada posición del lote. En este caso las dosis serán uniformes por zona de manejo.

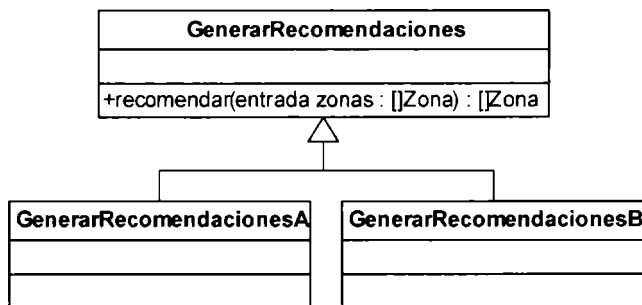


Imagen 14: Generación de recomendaciones

La posibilidad de obtener distintas recomendaciones de aplicación para la variedad de herramientas existentes se logra gracias a la capacidad de intercambiar los Generadores de Recomendaciones sin impactar en las opciones elegidas para todo el análisis previo.

4.6. Estrategia completa de análisis

La estrategia de análisis completa también posee características del patrón Abstract Factory que a continuación se detalla.

Patrón Abstract Factory

Propósito:

Provee una interfaz para la creación de familias de objetos relacionados o dependientes sin especificar su clase concreta.

Estructura:

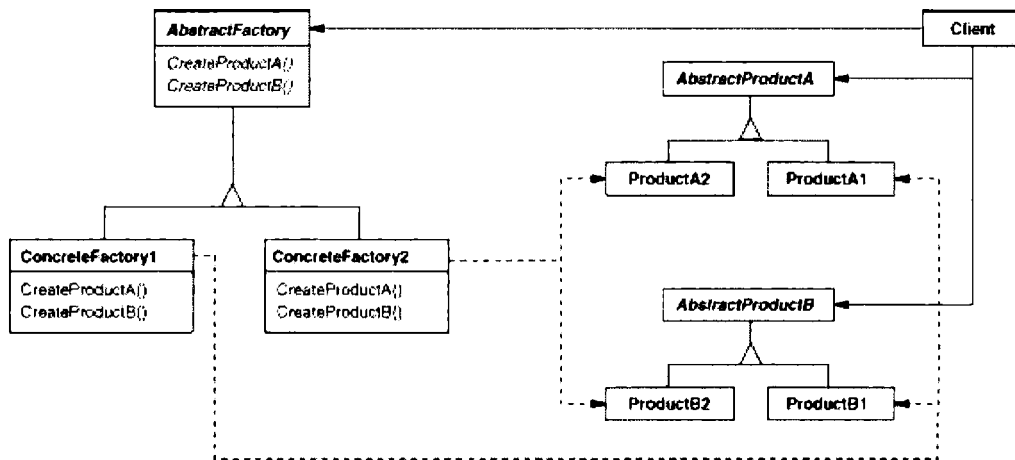
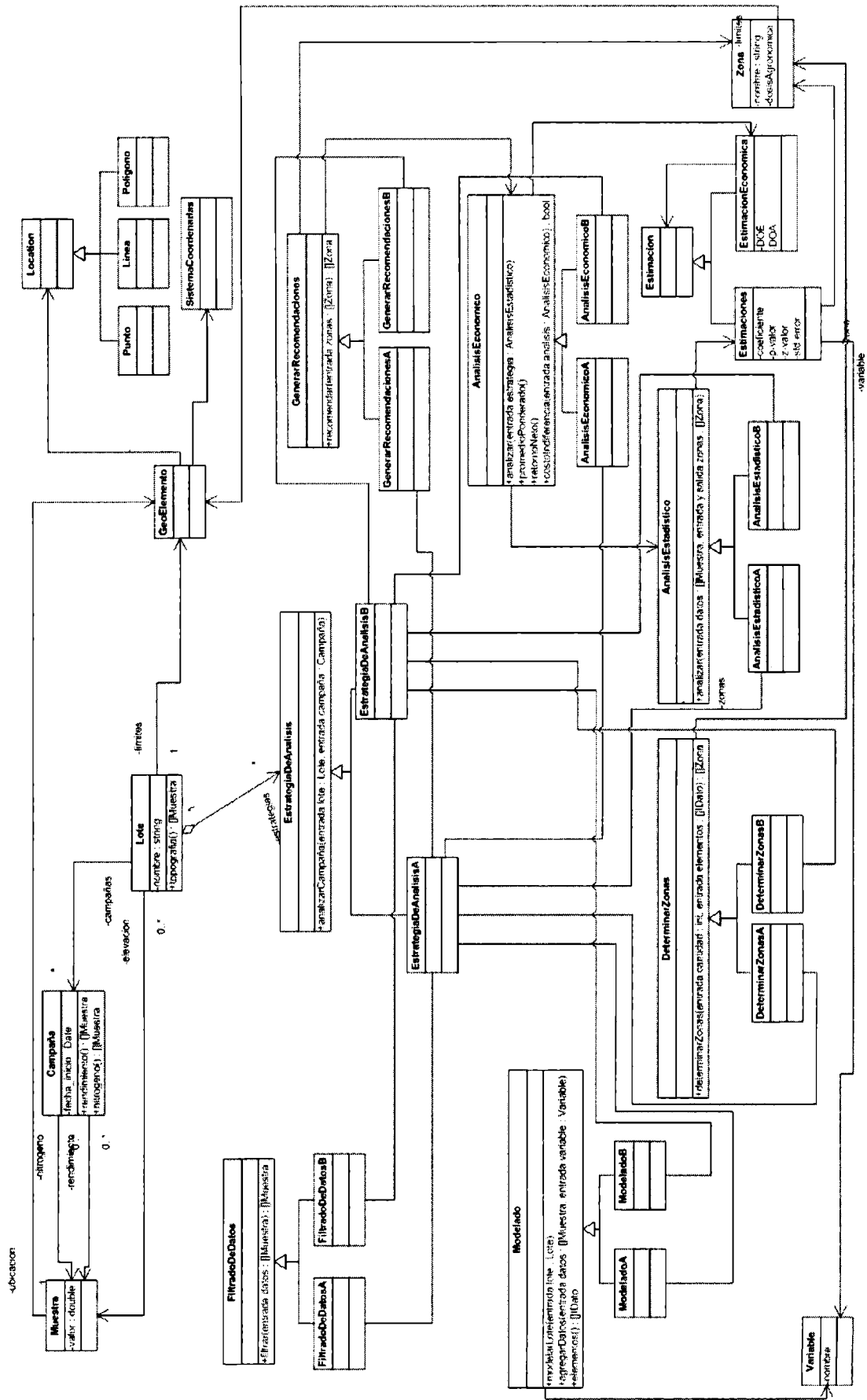


Imagen 15: Estructura del patrón AbstractFactory

- **AbstractFactory:** Declara una interfaz para la creación de objetos abstractos.
- **ConcreteFactory:** Implementa las operaciones de creación de objetos concretos.
- **AbstractProduct:** Declara una interfaz para un tipo de objeto.
- **ConcreteProduct:** Define un objeto concreto e implementa la interfaz declarada por el AbstractProduct.
- **Client:** utiliza las interfaces declaradas por el AbstractFactory y el AbstractProduct.

4.7. Arquitectura completa:

Finalmente se presenta la arquitectura completa de análisis donde se compone el modelo abstracto a partir de los modelos propuestos para cada una de los elementos identificados.



Capítulo 5: Instancia de la arquitectura

En este capítulo se presenta una instancia de la arquitectura propuesta que contempla el proceso de análisis llevado a cabo por Rodolfo Bongiovanni en su tesis doctoral utilizando técnicas de econometría espacial. Se toma como base la arquitectura propuesta en el capítulo anterior.

5.1. El Lote y las Campañas

Los lotes y las campañas se representan de manera similar a la abstracta, la topografía se representa en el lote ya que no varía entre campañas. Del Lote se conocen las campañas pasadas y sus límites geográficos representados por un GeoElemento. Las campañas son explotaciones que se inician en una fecha sobre un lote y contienen información muestreada de la aplicación de Nitrógeno y sobre el rendimiento logrado.

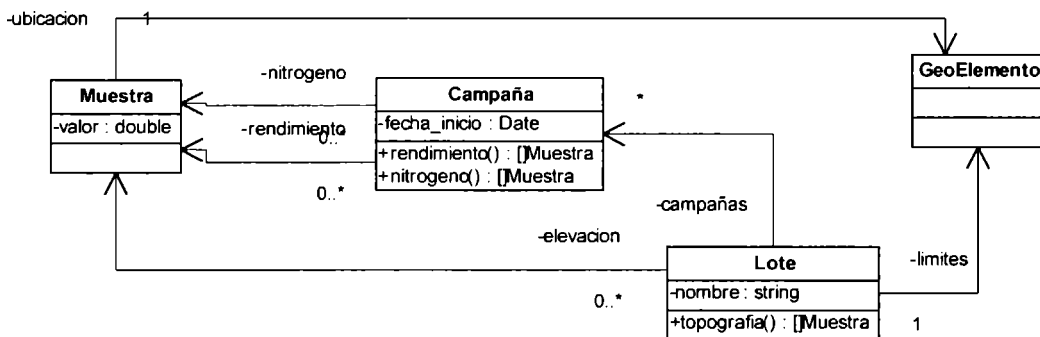


Imagen 1: Un Lote y sus Campañas

5.2. Estrategias de Análisis:

La estrategia de análisis abstracta se extiende para contemplar el proceso de análisis llevado a cabo por Rodolfo Bongiovanni en su tesis doctoral. La clase `AnalisisBongiovanni` es la base de la arquitectura y define tanto la estrategia como la forma en la que se generan las instancias cumpliendo su rol de factoría como se describió en el capítulo anterior, ya que se compondrá finalmente de cada una de las instancias de los procesos invocados; como filtrado, modelado, etc., configurados para el análisis en particular.

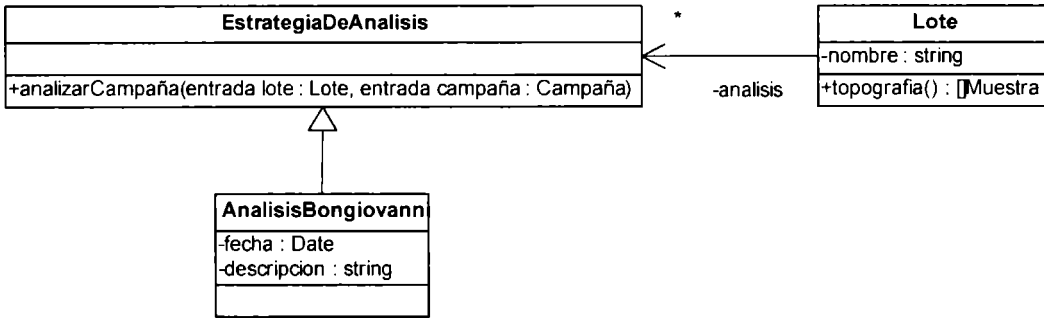


Imagen 2: Instancia del Patrón Strategy para el Análisis

5.3. Los Procesos

Cada uno de los modelo de proceso que se describirán cubrirán las necesidades planteadas para llevar a cabo el análisis particular planteado en este trabajo con técnicas de econometría espacial descriptas en el capítulo 3.

5.3.1. Filtrado de Datos

El filtrado de datos permite aplicar diversos criterios para seleccionar las muestras que serán descartadas del análisis. Los filtros pueden ser varios y se podría interactuar con otras herramientas para definir predicados sobre las muestras.

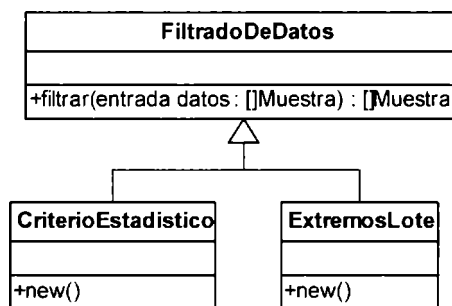


Imagen 3: Filtrado de datos

La interfaz propuesta en la superclase permite de una manera sencilla la composición de varios de estos procesos en forma secuencial.

5.3.2. Modelado

En la etapa de modelado se divide la superficie del lote en una partición formada por cuadrículas uniformes para generar elementos discretos y posibilitar el análisis, para cada una de estas cuadrículas se conoce los promedios de las muestras contenidas que pasaran a ser una Variable más que se evaluará en los siguientes procesos.

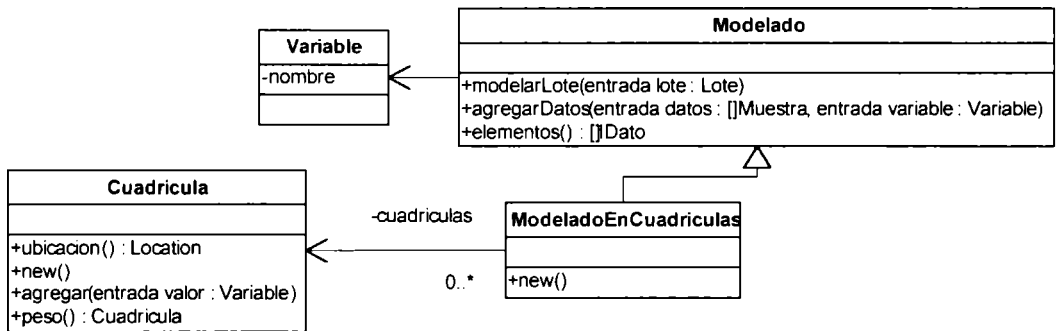


Imagen 4: Modelado

Este paso podría ser reemplazado por otras técnicas para generar otros elementos discretos necesarios para el análisis, por ejemplo el agregar a la colección de muestras más dispersas las muestras más densas en forma de promedios, etc.

5.3.3. Determinación de Zonas de Manejo

La intención de este modelo es agrupar los elementos discretos generados en el proceso anterior, que en el caso particular del análisis cubierto en este trabajo es una partición de la superficie del lote en cuadrículas, en zonas de manejo. La posibilidad de generar distintas variantes de agrupamiento con respecto al número final de zonas deseadas como se describe en el capítulo 3.3 permite comparar cada una de las opciones elegidas entre sí mediante indicadores acordes.

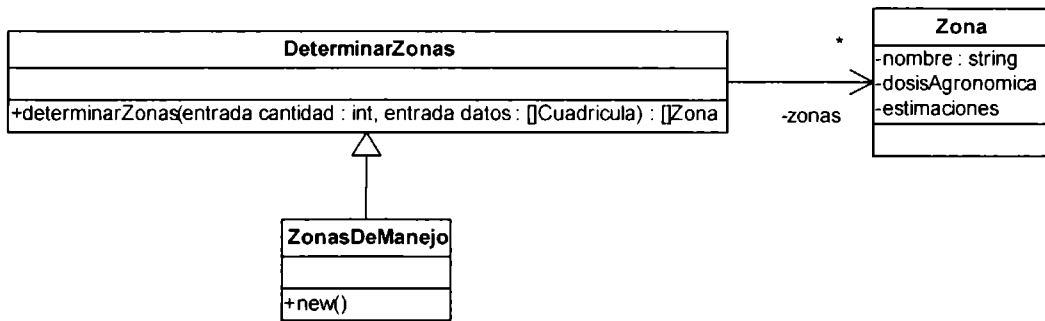


Imagen 5: Determinación de zonas de manejo

El agrupar las cuadrículas uniformes en zonas de manejo permite simplificar los mapas de aplicación ya que se estiman dosis uniformes para toda la zona en cuestión.

5.3.4. Análisis Estadístico

La opción de realizar múltiples análisis estadísticos basados en técnicas distintas provee herramientas para fundamentar con mayores bases las decisiones a tomar en futuras campañas. Cada uno de estos análisis son especializaciones del Análisis Estadístico.

Los dos análisis modelados en esta instancia son OLS tradicional que no incorpora la relación espacial entre las variables discretas obtenidas por los procesos anteriores y SAR que es el modelo espacial autorregresivo que incorpora explícitamente esta relación.

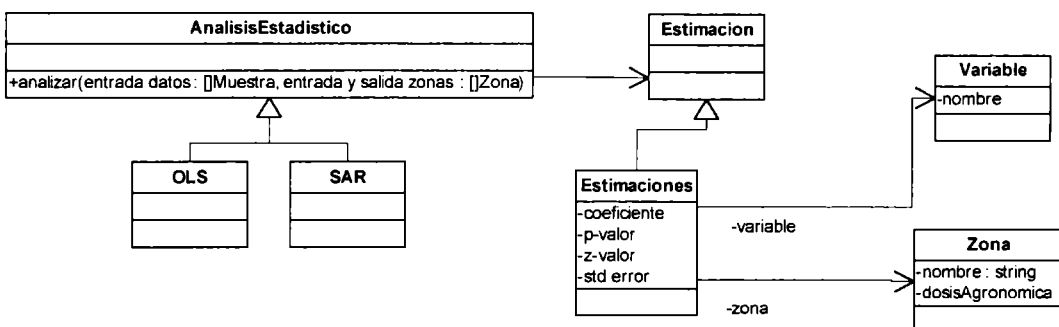


Imagen 6: Análisis estadístico

El proceso de Análisis Estadístico genera estimaciones para cada Zona de Manejo en la que se ha agrupado los elementos discretos, en este caso las cuadrículas uniformes,

para las distintas Variables que se incorporaron al análisis en la etapa de Modelado, como la dosis de Nitrógeno por zona y los datos del monitor de rendimiento filtrado y promediado, reflejando las necesidades descritas en el capítulo 3.

Los indicadores incorporados en cada una de las estimaciones son usados para medir la confianza que se puede tener en los estimadores obtenidos posibilitando la comparación entre los procesos llevados a cabo y permitiendo reducir el nivel de incertidumbre.

5.3.5. Análisis Económico

Tomando como base los Estimadores generados en el Análisis Estadístico recibido como entrada se llevan a cabo los análisis económicos de la explotación que nos dará información imprescindible a la hora de fundamentar las decisiones que se tomarán.

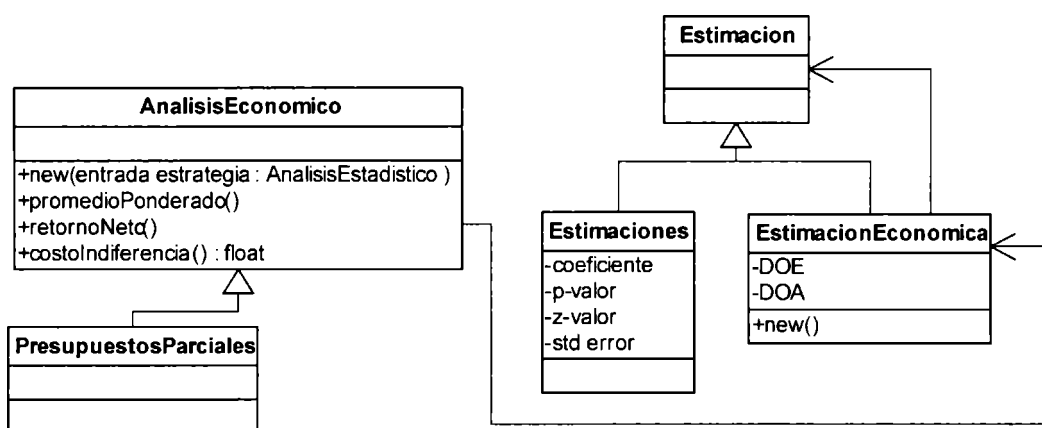


Imagen 7: Análisis económico

Para estimaciones particulares, como el Nitrógeno, se determinaran valores que se utilizaran en el Análisis Económico como la Dosis Óptima Económica o la Dosis Óptima Agronómica. A partir de estos valores se determinan los costos de producción y de indiferencia y se utilizan como soporte en la toma de decisiones.

5.3.6. Generación de Recomendaciones

Los mapas de recomendaciones son generados para programar las máquinas que realizarán las tareas de dosificación variable. Es posible subclasificar o adaptar los

generadores de mapas de aplicación para cubrir los diferentes formatos de entrada que soportan las máquinas construidas por distintos fabricantes.

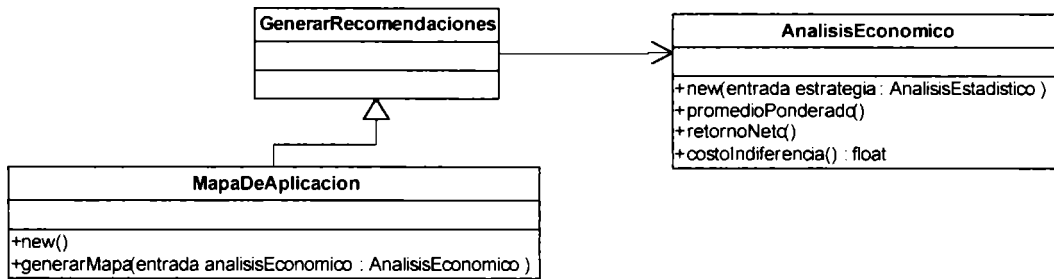
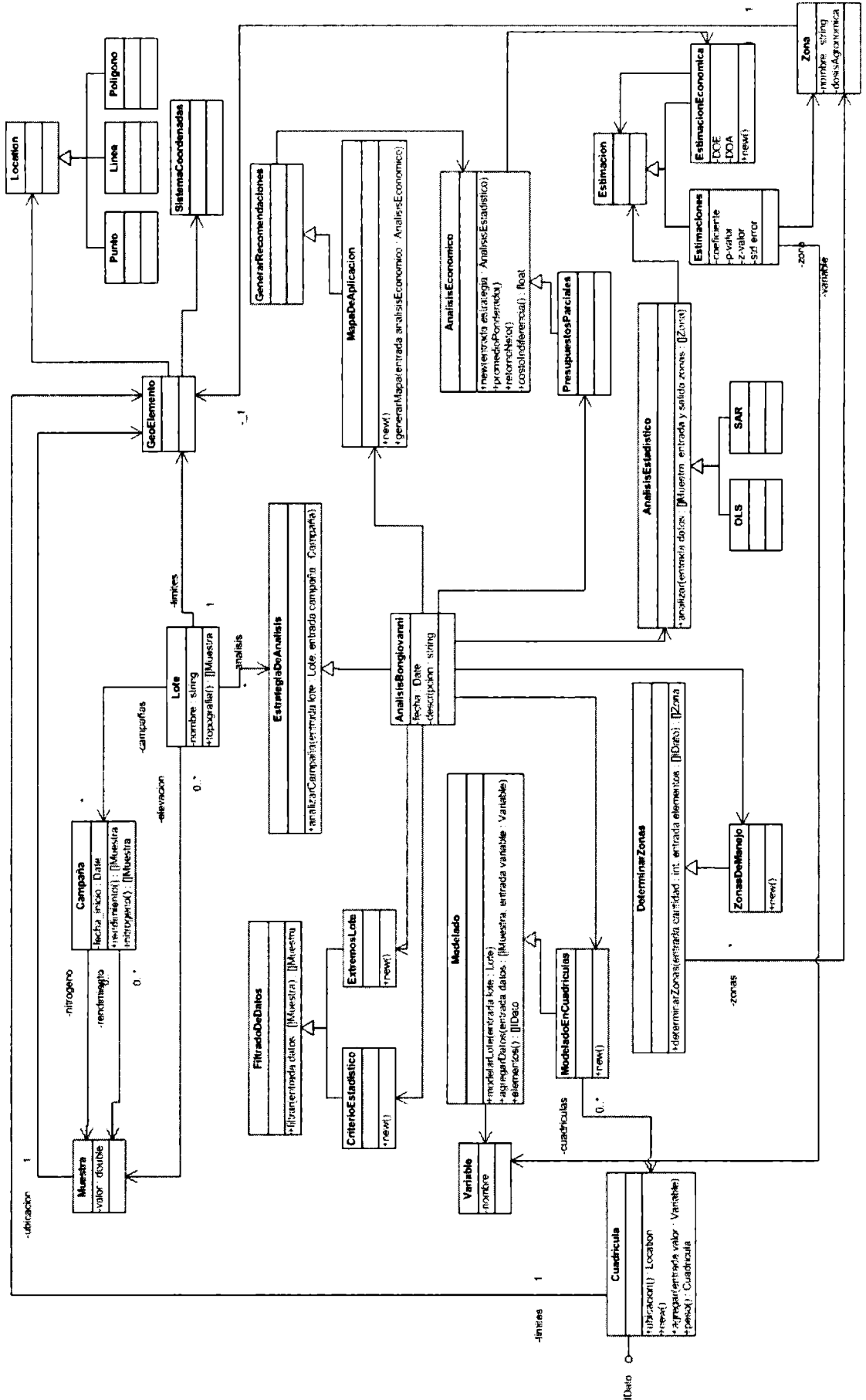


Imagen 8: Generación de recomendaciones

Se toma como base para determinar la dosis por posición la Dosis Óptima Económica obtenida por el análisis económico recibido como parámetro a la hora de generar el mapa.

5.4. Arquitectura completa



5.5. Diagrama de secuencia del análisis utilizando econometría espacial

Capítulo 6: Conclusiones

En este capítulo se presentan las conclusiones del trabajo realizado y los beneficios que se identifican en la aplicación de la arquitectura propuesta, de manera abstracta primero y como instancia del análisis llevado a cabo por Rodolfo Bongiovanni luego, en procesos de análisis de datos en la Agricultura de Precisión.

6.1. Conclusiones

Se presentó una arquitectura abstracta que contempla las características generales de los análisis de datos en la Agricultura de Precisión existentes revisados en la bibliografía, brindando un ambiente dinámico y fácilmente extensible a futuras estrategias de análisis integradas por nuevos procesos de tratamiento, modelado y análisis de datos.

También se propone una instancia del modelo abstracto en una arquitectura que resuelve particularmente la estrategia llevada a cabo por Rodolfo Bongiovanni en su tesis doctoral, la misma aplica los principios de la econometría espacial para contemplar la relación en el espacio de los datos recogidos como muestras.

Los siguientes beneficios se identifican en el presente trabajo:

- En el análisis de los datos del monitor de rendimiento se utilizan diversos criterios, herramientas y formatos para llevar a cabo el trabajo, volviendo difícil y falto de claridad el seguimiento de proyectos de análisis y la extracción de conclusiones de los mismos. La propuesta integra todos los pasos del análisis en un mismo sistema brindando claridad en el seguimiento y la revisión del proceso llevado a cabo sobre las muestras.
- Las distintas estrategias existentes de menor y mayor complejidad son llevadas a cabo con distintos criterios dificultando la comparación entre ellas. Esta arquitectura permite modelar las diferentes estrategias y realizar comparaciones entre ellas a pesar de ser compuestas por procesos diferentes ya que provee un marco similar de desarrollo y unifica criterios a la hora de medir los resultados.
- Para una misma estrategia es posible seleccionar distintas configuraciones de procesamiento, como filtros aplicados a las muestras, determinación de diferentes números de zonas de manejo, etc. El modelo propuesto resuelve

esto gracias a los beneficios de la Metodología Orientada a Objetos descripta en el Capítulo 2 que gracias a la herencia, el encapsulamiento y el polimorfismo, brinda un ambiente fácilmente configurable y extensible a nuevas formas de procesamiento. Todo esto, al pertenecer a la misma jerarquía permite realizar comparaciones entre los caminos elegidos para dar un soporte más confiable a las decisiones que serán tomadas en futuras campañas.

- En un mismo análisis pueden participar varias herramientas volviendo indispensable la transformación de los datos de un formato a otro para la integración de las mismas. La arquitectura descripta en el trabajo soluciona esto ya que todo el proceso se lleva a cabo en el mismo sistema sin resultar necesario la integración con herramientas externas salvo que el proceso lo requiera, siendo responsabilidad de este modulo en cuestión la transformación, exportación e importación a formatos externos.

En este trabajo se omite la resolución interna de cada modulo de proceso, esta etapa queda como parte de trabajo futuros.

6.2. Trabajos futuros.

Los trabajos futuros están orientados a continuar el presente trabajo de grado y propone finalmente algunas opciones específicas para mejorar las características de extensibilidad de la arquitectura propuesta:

- Implementación del comportamiento de cada una de las clases de la instancia de la arquitectura propuesta en el capítulo 5.
- Extensión de la jerarquía de estrategias de análisis para contemplar otras existentes o de futura creación, esto ampliará el conjunto de posibles usuarios del sistema al incluir su metodología de trabajo en las posibles instancias.
- Extensión de las colecciones de procesos disponibles para llevar a cabo las estrategias implementadas de manera de incluir nuevos procesos de análisis no contemplados en el presente trabajo para cubrir la demanda de usuarios específicos.
- El rol de factoría de las estrategias de análisis puede verse beneficiado por técnicas de Inversión de Control (IOC) para permitir ser configuradas

externamente mediante archivos XML. Esto significa definir externamente la forma en que se componen los procesos en archivos que indican que módulos se utilizaran y de que manera se ensamblan para definir la estrategia de análisis.

- Cada uno de los procesos de análisis pueden resolverse mediante la invocación a servicios remotos. Mediante llamadas a servicios Web con interfaces estándar y definidas por organizaciones externas se puede elevar el volumen de procesos disponibles para llevar a cabo el análisis deseado.

Capítulo 7: Referencias Bibliográficas

- Anselin L., Bongiovanni R. y Lowenber-DeBoer J. (2004, Agosto). A Spatial Econometric Approach to the Economics of Site-Specific Nitrogen Management in Corn Production. *Amer. J. Agr. Econ* 86(3): 675-687
- Anselin, L. (1988). *Spatial econometrics: methods and model*. (Kluwer Academic Publishers: London)
- Anselin, L. (2001). Spatial Effects in Econometric Practice in Environmental and Resource Economics, *American Journal of Agricultural Economics* 83, 705-710.
- Baños A. y Goenaga T. (2003, Junio). Metodología para la evaluación económica de un proyecto de Agricultura de Precisión. www.ideared.org/doc/investigacion_agricultura_de_precision.pdf. Acceso: 19 de abril de 2006.
- Bongiovanni, R. (2002a). A Spatial Econometric Approach to the Economics of Site-Specific Nitrogen Management in Corn Production. Ph.D. Thesis, *Agricultural Economics*, Purdue Univ. 306 pp. <http://www1.lib.purdue.edu/>.
- Bongiovanni R. (2002b). Método para Establecer el Valor de la Agricultura de Precisión: Respuesta del Maíz al Nitrógeno en Ensayos a Campo. Trabajo preparado para el 4^{to} Curso de Agricultura de Precisión, INTA-Manfredi, 24-25 de Julio de 2002.
- Bongiovanni, R. y J. Lowenberg-DeBoer. (2006). Capítulo 9. VIABILIDAD ECONÓMICA en el libro. *Agricultura de Precisión en el Cono Sur de América*. Bongiovanni, Mantovani, Best, y Roel (editores). Publicado por PROCISUR (en imprenta).
- Bongiovanni, R., y Lowenberg-DeBoer, J. (2001). "Agricultura de Precisión y Sustentabilidad". Trabajo presentado como conferencista invitado en el VII^o Congreso de Maíz, Pergamino, Argentina, Noviembre 7-9, 2001.
- Bragachini M., Bongiovanni R., Méndez A. y Scaramuzza F. (2005). FERTILIZACION Y DENSIDAD DE SIEMBRA VARIABLE. En: <http://www.agriculturadeprecision.org/mansit/FertilizacionyDensidadSiembraVariable.htm>. Acceso: 19 de abril de 2006
- Bullock D. G., Bullock D. S., Bollero G. y Ruffo M. (2003, septiembre). Agricultura de precisión: Necesitamos algo más que tecnología. En www.inta.gov.ar/balcarce/info/documentos/agric/sd/bullock.htm Acceso: 19 de abril de 2006
- Calp D. (1999, Agosto). COMPUTACION AGROPECUARIA. PRIMEROS PASOS CON GIS. En: www.aacrea.org.ar/economia/articulos/pdf/art226b.pdf. Acceso: 30 de abril de 2006

- Casady W., Pfoest D., Ellis C. y Shannon K. (1998). Precision Agriculture: Yield Monitors. En mpac.missouri.edu/pubs/wq0451.pdf Acceso: 19 de abril de 2006
- Dobermann A., Blackmore S., Cook S. E. y Adamchuk V. I. (2004). Precision Farming: Challenges and Future Directions. Proceedings of the 4th International Crop Science Congress, Brisbane, Australia.
- Emmen, D. A. (2004). La agricultura de precisión: una alternativa para optimizar los sistemas de producción. *Invest. Pens. Crit*, 2, 68-74.
- Erickson B. (2005, Agosto). Step-by-Step Yield Monitor Data Analysis. En Site Specific Management Center Newsletter, Purdue University. En www.agriculture.purdue.edu/ssmc/frames/ssmc_YMA_8_05.pdf Acceso: 19 de abril de 2006.
- Gamma E., Helm R., Jonson R. y Vlissides J. (1995). Design Patterns.
- Goddard T., Kryzanowski L., Cannon K., Izaurrealde C. y Martin T. (1995). Potential for Integrated GIS-Agriculture Models for Precision Farming Systems. En www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papers/goddard_tom/960119.html Acceso: 19 de abril de 2006.
- Gordillo, S. (Octubre, 1998) "Modelización de Campos Continuos en Sistemas de Información Geográfica". Tesis de Maestría. Director: Gabriel Baum
- Gregoret, M.C; Dardanelli, J; Bongiovanni, R.G.; J; Diaz Zorita, M. (2005). ANÁLISIS DE LA RESPUESTA SITIO-ESPECÍFICA AL NITRÓGENO EN MAÍZ. PARTE I: CARACTERIZACIÓN DE AMBIENTES. Actas del VIII Congreso Nacional de Maíz 2005. Rosario, 16-18 Noviembre 2005. pag.137-140.
- Griffin T. W., Brown J. P. y Lowenberg-DeBoer J. (2005). YIELD MONITOR DATA ANALYSIS: DATA ACQUISITION, MANAGEMENT, AND ANALYSIS PROTOCOL. www.agriculture.purdue.edu/ssmc/publications/YieldDataAnalysis.pdf Acceso: 19 de abril de 2006.
- Hatfield J. 2000. Precision Agriculture and Environmental Quality: Challenges for Research and Education. National Soil Tilth Laboratory, Agricultural Research Service, USDA, Ames, Iowa. Documento disponible en <http://www.arborday.org>. Consultado el 12/12/04.
- Lachapelle, G., Cannon, M.E., Gehue, H., Goddard, T. and Penney, D. (1994) GPS system integration and field approaches in precision farming. *Navigation* 41(3): 323-335.
- Lambert M. D., Lowenberg-DeBoer J. y Bongiovanni R. 2004. A Comparison of Four Spatial Regression Models for Yield Monitor Data: A Case Study from Argentina *Precision Agriculture*, 5, 579-600, 2004. 2005 Kluwer Academic Publishers.

- Longley P. A., Goodchild M. F., Maguire D. J. y Rhind D. W. (1999). Geographical Information Systems, Principles, Techniques, Applications and Management, 2 Volume Set, 2nd Edition.(Wiley): 13-14
- Lowenberg-DeBoer, J. and Swinton, S. (1997). "Economics of Site-Specific Management in Agronomic Crops," Chapter 16. In: Pierce, F., and E. Sadler, eds. The State of Site-Specific Management for Agriculture, (ASA-CSSA-SSSA, Madison, Wisconsin, 1997), pp. 369-396.
- Luchiari, A. 2000. Strategies For Establishing Management Zones For Site Specific Nutrient Management. Proc. 5th Intern. Conference on precisión Agriculture. Precision Agriculture Center, ASA, CSSA and SSSA, MN.
- Melchiori R. J. M., García F. O. y Echeverría H. (2000). VARIABILIDAD ESPACIAL EN ALGUNAS PROPIEDADES DEL SUELO: II MANEJO DEL N POR SITIO ESPECÍFICO EN EL CULTIVO DE TRIGO. En www.agriculturadeprecision.org/mansit/VariabilidadEspacial2.htm Acceso: 19 de abril de 2006
- Mizzou-ARS (University of Missouri-Columbia & Agricultural Research Service of the United States Department of Agriculture). 2000. Management Zone Analyst Version 1.0.1. <http://www.ars.usda.gov/services/software/download.htm?softwareid=24> (May 23, 2006)
- Roberts T. L. (2000). Manejo sitio específico de Nutrientes – Avances en Aplicaciones con Dosis Variable. En <http://www.agriculturadeprecision.org/mansit/ManejoNutrientesSitioEspecifico.htm> Acceso: 30 de abril de 2006
- Rumbaugh J., Blaha M., Premerlani W., Eddy F. y Lorenzen W. (1991). Object-Oriented Modeling and Design. Englewood Cliffs, New Jersey: Prentice Hall
- Saraiva A. M., Massola A. M. A. y Cugnasca C. E. (1998). An Object Model for Information System for Precision Agriculture. Proceedings of the 4th International Conference on Precision Agriculture. 1355-1366.
- Siegel S. (Julio, 2005). Introduction to OMG's Unified Modeling Language. En: http://www.omg.org/gettingstarted/what_is_uml.htm Acceso: 8 de octubre de 2006
- von Martini A., Méndez A. y Bragachini M. (2003). Potencialidad de la información de los Mapas de Rendimiento en la Interpretación del Manejo de Factores de Rendimiento y su Respuesta Sitio Específica en el Cultivo de Soja. En: www.agriculturadeprecision.org/monrend/PotencialidadInformacionMapasRendimiento.pdf Acceso: 19 de abril de 2006.