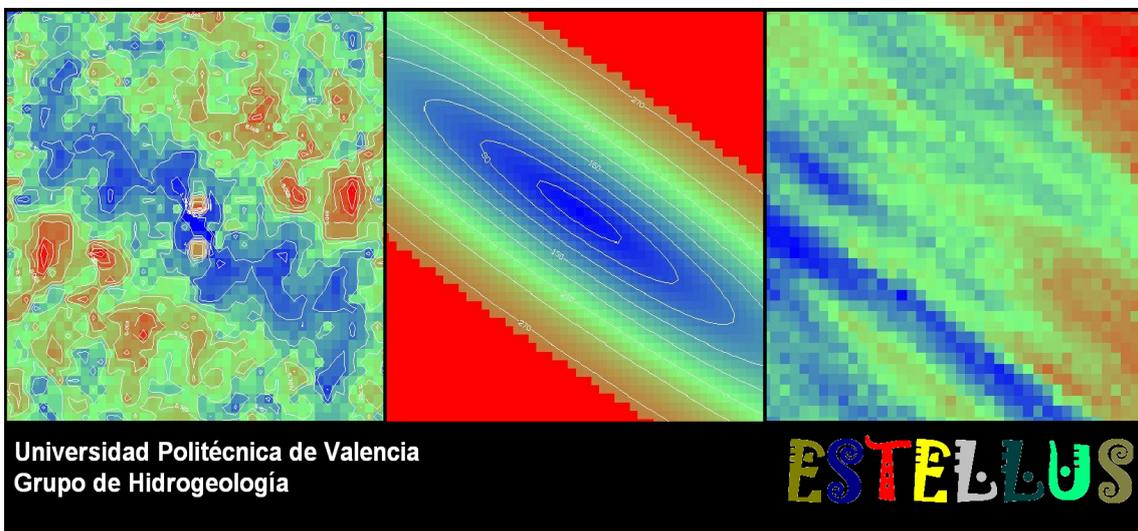


DESARROLLO DE UNA LIBRERÍA COMPUTACIONAL PARA LA ESTIMACIÓN ESPACIAL UTILIZANDO CONCEPTOS GEOGRÁFICOS Y PROGRAMACIÓN ORIENTADA A OBJETOS ESPACIALES



Trabajo de fin de máster presentado por:
Oscar David Álvarez Villa

Dirigido por:
Eduardo Cassiraga



grupo
de **HIDROGEOLOGÍA**

**DESARROLLO DE UNA LIBRERÍA COMPUTACIONAL
PARA LA ESTIMACIÓN ESPACIAL UTILIZANDO
CONCEPTOS GEOGRÁFICOS Y PROGRAMACIÓN
ORIENTADA A OBJETOS ESPACIALES**

Trabajo de fin de máster presentado por:
Oscar David Álvarez Villa

Dirigido por:
Eduardo Cassiraga

**Grupo de Hidrogeología
Departamento de Ingeniería Hidráulica y Medio Ambiente
Universidad Politécnica de Valencia
Diciembre de 2009**

Dirección del autor

Oscar David Álvarez Villa

Grupo de Hidrogeología
Departamento de Ingeniería Hidráulica y Medio Ambiente
Universidad Politécnica de Valencia
Camino de Vera, S/N
46022 Valencia, España

e-mail: odalvare@gmail.com, osdaalvi@gmail.com

tel: + 57 4 250 03 77

móvil: + 57 300 8058864 (Colombia); + 34 699 754 152 (España)

A mi madre....

*Madre, llévame a la cama.
Madre, llévame a la cama,
que no me tengo de pie.
Ven, hijo, Dios te bendiga
y no te dejes caer.*

*No te vayas de mi lado,
cántame el cantar aquél.
Me lo cantaba mi madre;
de mocita lo olvidé,
cuando te apreté a mis pechos
contigo lo recordé.*

*¿Qué dice el cantar, mi madre,
qué dice el cantar aquél?
No dice, hijo mío, reza,
reza palabras de miel;
reza palabras de ensueño
que nada dicen sin él.*

*¿Estás aquí, madre mía?
porque no te logro ver....
Estoy aquí, con tu sueño;
duerme, hijo mío, con fe.*

Madre, llévame a la cama
Rafael Alberti

Agradecimientos

Deseo expresar mi agradecimiento al Prof. Eduardo Cassiraga, director de este trabajo, por todo el apoyo que me ha brindado durante este tiempo y por estar siempre dispuesto a colaborar en cualquier problema, sea de índole académico como personal. Igualmente gracias a los personas pertenecientes al Grupo de Hidrogeología del Departamento de Ingeniería Hidráulica y Medio Ambiente de la Universidad Politécnica de Valencia por todo el apoyo recibido: a los profesores Jaime Gómez Hernández y Andrés Sahuquillo; a Francisco Morrillas, Liangping Li, Haiyan Zhou, Teresa Martínez y Pilar Mejía. Un enorme sentimiento de gratitud para los profesores Jaime Ignacio Vélez y Germán Poveda del Postgrado en Aprovechamiento de Recursos Hidráulicos de la Universidad Nacional de Colombia (Sede Medellín), quienes me ayudaron a desarrollar parte del trabajo y me han apoyado desde hace tanto tiempo. De nuevo y por tercera vez, gracias a Olver Hernández y Felipe Quintero, mis amigos, maestros de Java y compañeros de aventura en Europa.

No tengo como agradecer a mi familia por haberme dado todo su apoyo y por motivarme para seguir adelante, especialmente en los momentos más difíciles. Muchas gracias a mi padre Arturo, mi hermana Angela, mi tía Marta, mis primos Carlos, Astrid, Paula, Bibiana y Andrés. Millones de gracias a Carolina María, tú eres mi inspiración.

Muchas gracias a los amigos que me ayudaron a familiarizarme con mi nuevo estilo de vida en España y me dieron consejos para que todo me resultara más sencillo: Diana Quevedo, Matías Peredo, Camilo Múnera y Verónica Pavón. También mil gracias para todos mis compañeros de máster en Valencia, especialmente para quienes me han brindado su apoyo, motivación y con quienes he compartido muy buenos momentos: Carlos Castro, Eliana Molina, Carolina Rubio, Lorena Sánchez, Luis Chaparro, Luis García, Elisa Martínez y Claudia Garzón. Gracias también a los amigos que se preocupan por mi desde la distancia: Jorge Esteban Alarcón, Carlos *El Sucio* Restrepo, Joany *El Dirty* Sánchez, Jorge *El Ácaro* Hernández, Cristian *Demonmalfas* Ortiz, Oscar *El Salsa* Estrada, Juliana Ospina, Marcela Cadavid, Juan David Correa, Juan Fernando Patiño, Sandra López, Oscar Ardila, Alejandro Cadena, Diana Restrepo, Lina Marín, Paola Roldán, Paola Arias, Yuley Cardona, Carolina Ortiz y Oscar Rueda. También, gracias a todas las personas que estuvieron a mi lado en la muerte de mi madre, su compañía me hizo sentir mejor en los momentos de dolor.

Finalmente, gracias mamá por brindarme siempre tu amor, apoyo, comprensión,..., gracias por todo lo que me diste y enseñaste. Siempre vas a estar presente en mi corazón.

Resumen

Este trabajo final de máster trata acerca de la construcción de una librería de estimación geoestadística que permita la consideración explícita de la localización espacial de los datos de entrada y los resultados obtenidos mediante su georeferenciación a partir de sistemas de coordenadas geográficos o planos estándar. Para lograr dicho objetivo las entidades espaciales susceptibles de integrar en un análisis estimativo se conceptualizan como objetos computacionales que tienen asociados diferentes tipos de propiedades medibles, observables y/o calificables. A dichos objetos se les puede asociar ciertas localizaciones referidas a un origen de coordenadas geográfico. Así, desde el punto de vista computacional se utiliza la programación orientada a objetos con el fin de modelar entidades computacionales que representen a las reales.

Inicialmente, se trata el desarrollo de los objetos computacionales que manipulan la información de entrada y de salida a los algoritmos implementados. En general, dado que la información utilizada en la geoestadística se basa en datos (mediciones) localizados puntualmente o en información densamente muestreada, los objetos computacionales espaciales que se modelan son entidades puntuales y campos en formato raster. La creación de los *objetos* computacionales se implementa en *clases* individuales que hacen posible la definición de su propiedades, su localización y su sistema de georeferenciación en los *constructores* asociados. Dado que el *objeto* debe representar características espaciales definidas, para una misma entidad espacial (por ejemplo localizaciones puntuales de mediciones) se generan diferentes tipos de objetos de acuerdo con las dimensiones del análisis: 2D o 3D. Para la manipulación y operación con los objetos modelados computacionalmente se generan los respectivos *métodos* de clase, que pueden ser *privados* o *públicos*. Los métodos públicos pueden ser utilizados efectivamente para nuevo desarrollo del software, por lo cual a lo largo del trabajo se hace referencia exclusivamente a ellos.

El aspecto siguiente que se trata es la implementación de los algoritmos de análisis de información: el análisis estadístico descriptivo de la información disponible y el análisis estructural con el fin de proponer un modelo que explique el comportamiento espacial del fenómeno estudiado. Ambos procedimientos deben realizarse como paso previo a la ejecución de cualquier algoritmo geoestadístico de estimación. La implementación de las herramientas computacionales para el análisis de información se realiza en *clases* especiales, cuyos *parámetros* fundamentales de construcción son *objetos* geográficos. Los *métodos* de clase implementan los algoritmos de cálculo propiamente dichos. En el caso del análisis estadístico se pueden realizar descripciones univariadas o bivariadas, mientras que las clases

desarrolladas para el análisis estructural permiten realizar estimaciones de variogramas experimentales y ajustes automáticos de modelos de variograma.

El aspecto más importante que trata este trabajo es la implementación informática de los diferentes algoritmos de estimación espacial. Con respecto a ello se realiza una evaluación crítica de las diferentes metodologías existentes para la estimación espacial de variables que se encuentran actualmente en la literatura. Las metodologías clásicas para la estimación espacial como los polígonos de Thiessen o el método del inverso cuadrático de la distancia, no consideran que la variable a estimar es aleatoria, por lo cual no puede cuantificarse la confiabilidad asociada con las estimaciones. Por el contrario, los algoritmos geoestadísticos consideran explícitamente la aleatoriedad de los fenómenos representados y, en virtud de lo anterior, permiten cuantificar la varianza relacionada con la estimación espacial. Igualmente, se presenta una revisión de diferentes librerías de estimaciones espaciales existentes y sus características se analizan en profundidad para tratar de incorporar sus ventajas en la librería de algoritmos desarrollada en este trabajo.

Así, para cada metodología de estimación se implementa una clase, es decir, se crea un objeto computacional que puede realizar procedimientos de estimación. Los *objetos* de estimación están caracterizados básicamente por la información utilizada, representada mediante un objeto computacional geográfico, y por el modelo de variación espacial construido para el atributo que se desea estimar. Como resultado los métodos arrojan objetos espaciales raster que representan un campo distribuido de los valores estimados en un área de estudio. Se implementan algunas metodologías deterministas: los polígonos de Thiessen y la estimación con ponderaciones basadas en el inverso de la distancia a un exponente; junto con métodos geoestadísticos: krigado ordinario, krigado simple, krigado con deriva externa, cokrigado simple, cokrigado ordinario, cokrigado colocalizado bajo un modelo de Markov.

Finalmente, para posibilitar el interacción de un usuario de nivel intermedio con los algoritmos implementados, se muestra una interfaz gráfica de usuario que gestiona y administra los procesos computacionales relacionados con la ejecución de los diferentes análisis integrados a la librería y permite al usuario realizar el ingreso de datos y la extracción, visualización y análisis de los resultados. El software (librerías e interfaz) desarrollado se denomina ESTELLUS y está escrito completamente en lenguaje de programación orientado a objetos **Java** para posibilitar la implementación de la conceptualización propuesta. Además de lo anterior, **Java** hace posible el uso de una gran cantidad de librerías de acceso libre existentes, algunas de las cuales fueron adaptadas y utilizadas para propósitos de visualización de resultados y análisis de datos. ESTELLUS pueden ser usado efectivamente en estudios de diferente tipo: modelación hidrogeológica, análisis hidrológicos, estudios ambientales, análisis meteorológicos, estudios de la ciencia del suelo, análisis de riesgo, prospección minera, etc. Mediante la aplicación del software para la resolución de un problema práctico sencillo en hidrogeología se demuestra su versatilidad para el análisis y visualización de información geográfica y sus amplias posibilidades de expansión futura en lo concerniente con la estimación y simulación estocástica.

Abstract

This master thesis deals with the implementation of a geostatistical library for spatial estimation that allows to consider explicitly the spatial location of the data and the obtained results by using standard (cartesian or geographic) georeferencing coordinate systems. To achieve such goal, the spatial entities that are able to be used in an estimation analysis are conceptualized as computational *objects* or entities with some intrinsic *properties*: they possess different types of features that are able to be observed or qualified and they have associated some coordinates referred to a geographical reference system. Then, from a computational perspective, object oriented programming is used to model some computer objects that represent accurately such spatial entities.

Initially, the development of the computational objects to handle the input and output information to and from the estimation algorithms is presented. In general, because the information used in geostatistics is based in scattered-point data or is densely measured information, the modeled computational *objects* of spatial information are vectorial points and raster fields. The creation of the computational *objects* has been implemented in particular *classes*. Each *class* allows to define the spatial entity's properties, locations and georeferencing system inside of the associated *builders*. Because the *object* has to represent different spatial features for entities of the same type (for instance: scattered-pointed locations of measuring) different kinds of spatial objects are implemented according to the dimensions of the analysis: 2D or 3D. To handle and operate with the computationally modeled spatial objects, their associated *class methods* are developed. In object oriented programming, *class methods* can be *public* if they are able to be used outside of the container class or *private* if they are able to be only inside the container class. Only *public methods* can be effectively used to create extensions of the actual version of the software.

Another subject treated in this thesis is about the implementation of algorithms for the description and analysis of the information: statistical description of the data sample and structural analysis to propose models of variability that explain the spatial behavior of the analyzed phenomena. Both of them have to be made as the previous step of the execution of any estimation procedure. Information analysis algorithms have been implemented in *classes* having geographical *objects* as fundamental *parameters*. The *class's methods* implement the analysis algorithms. In the particular case of the statistical analysis, univariate or bivariate descriptions are able to be performed. In the case of the structural analysis, algorithms for experimental variogram estimation and automatic fitting of variogram models are able to be performed.

The most important topic treated in the thesis is the computational implementation of some algorithms for the spatial estimation. Regarding to them, a critical a thorough literature review about spatial estimation methodologies is presented. The classical methodologies to perform spatial estimations, for instance, Thiessen’s polygon or inverse distance interpolation, can not consider the randomness of the variable that has been estimated. In opposition, geostatistical algorithms can account the randomness of the represented phenomena and allow quantifying the reliability related with the spatial interpolation procedure. Also, a revision of the features of some computational libraries for spatial estimation is given. Their advantages and weakness are widely commented. In this work the main strengths of those libraries are adopted in the implementation of the estimation algorithms.

For each estimation methodology a *class* is implemented, it means that a computational *object* is created to be able to perform the estimation procedure. The estimation objects are characterized by two properties: the information used for estimation (spatial objects) and the model of spatial variability of the attribute. The results of the *class methods* are raster objects that represents a random field of the estimated variable over a given area. In this work, the most popular estimation algorithms have been implemented: Thiessen’s polygons, inverse distance interpolation, simple kriging, ordinary kriging, simple cokriging, ordinary cokriging, collocated cokriging and collocated cokriging using a Markov model.

Finally, with the porpoise of achieve the interaction of the user with the implemented algorithms, the development of a graphical interface is presented. This interface manage the computational procedures related with the execution of the different types of analyses programmed in the library and also allows to input the data and extract, view and study the results. ESTELLUS is the name given to the developed software (algorithm library plus user’s interface). It is written fully in *Java* object oriented programming language to be able to implement the conceptualization for the spatial objects. Also, *Java* makes posible to use a immense number of free access libraries, some of which were adapted to be used in ESTELLUS development, specially for data and results visualization and analysis of samples of information. ESTELLUS can be used effectively in different types of studies an researches: hydrogeological modeling, hydrological analysis, environmental studies, meteorological researches, soil science studies, risk analysis, minning propection, etc. ESTELLUS is used to solve a practical problem of spatial estimation showing its versatility for analysis and visualization of geographical information and their possibilities of future growth. It is expected to integrate the simulations algorithms to the library in the future along with more utilities for handling the data and results.

Resum

Este treball final de màster tracta sobre la construcció d'una llibreria d'estimació geostatística que permeta la consideració explícita de la localització espacial de les dades d'entrada i els resultats obtinguts per mitjà del seu georeferenciación a partir de sistemes de coordenades geogràfics o plans estàndard. Per a aconseguir dita objectiva lassa entitats espacials susceptibles d'integrar en una anàlisi estimatiu es conceptualitzen com a *objectes* computacionals que tenen associats diferents tipus de *propietats* mesurables, observables y/o qualificables. Als dits objectes se'ls pot associar certes localitzacions referides a un origen de coordenades geogràfic. Així, des del punt de vista computacional s'utilitza la programació orientada a objectes a fi de modelar entitats computacionals que representen les dites entitats espacials reals.

Inicialment, es tracta el desenrotllament dels *objectes* computacionals que manipulen la informació d'entrada i d'eixida als algoritmes implementats. En general, atés que la informació utilitzada en la geostatística es basa en dades (mediciones) localitzats puntualment o es troba densament muestreada, els *objectes* computacionals espacials que es modelen són entitats puntuals i camps format raster. La creació dels *objectes* computacionals s'implementen en classes individuals que fan possible la definició de les seues propietats, la seua localització i el seu sistema de georeferenciación en els *constructors* associats. Atés que l'*objecte* ha de representar característiques espacials definides, per a una mateixa entitat espacial (por exemple localitzacions puntuals de mediciones) es generen diferents tipus d'objectes d'acord amb les dimensions de l'anàlisi: 2D o 3D. Per a la manipulació i operació amb els objectes modelatges computacionalment es generen els respectius *mètodes de classe*, que poden ser *privats* o *públics*. Els *mètodes públics* poden ser utilitzats efectivament per a nou desenrotllament del programari, per la qual cosa al llarg del treball es fa referència exclusivament a ells.

L'aspecte següent que es tracta és la implementació dels algoritmes d'anàlisi d'informació: l'anàlisi estadístic descriptiu de les mostres de dades i l'anàlisi estructural a fi de proposar un model que explique el comportament espacial del fenomen estudiat. Ambdós procediments han de realitzar-se com a pas previ a l'execució de qualsevol algoritme geostatístico d'estimació. La implementació de les ferramentes computacionals per a l'anàlisi d'informació es realitza en *classes* especials, els *paràmetres* fonamentals de construcció de la qual són *objectes* de geogràfics. Els *mètodes de classe* implementen els algoritmes de càlcul pròpiament dites. En el cas de l'anàlisi estadística es poden realitzar descripcions univariadas o bivariadas, mentres que les *classes* desenrotllades per a

L’anàlisi estructural permeten realitzar estimacions de variogramas experimentals i ajustos automàtics de models de variograma.

L’aspecte més important que tracta este treball és la implementació informàtica dels diferents algoritmes d’estimació espacial. Respecte a això es realitza una avaluació crítica de les diferents metodologies existents per a l’estimació espacial de variables que es troben actualment en la literatura. Les metodologies clàssiques com els polígons de Thiessen o el mètode de l’invers quadràtic de la distància, no consideren que la variable a estimar siga aleatòria, per la qual cosa no pot quantificar-se la confiabilitat associada amb les estimacions. Al contrari, els algoritmes geoestadístics consideren explícitament l’aleatoriedad dels fenòmens representats i, en virtut de l’anterior, permeten quantificar la varianza relacionada amb l’estimació espacial. Igualment, es presenta una revisió de diferents llibreries d’estimacions espacials existents i les seues característiques s’anàlitzten en profunditat per a incorporar les seues fortaleces en la llibreria d’algoritmes desenrotllada en el present treball.

Així, per a cada metodologia d’estimació s’implementa una *classe*, és a dir, es crega un *objecte* computacional que pot realitzar procediments d’estimació. Els *objectes* d’estimació estan caracteritzats bàsicament per la informació usada, representada per mitjà d’un objecte computacional geogràfic, i pel model de variació espacial construït per a l’atribut que es desitja estimar. Com resultat els *mètodes* tiren objectes espacials raster que representen un camp distribuït de ho valors estimats en una àrea d’estudi. S’implementen algunes metodologies deterministes: els polígons de Thiessen i l’estimació amb ponderacions basades en l’invers de la distància a un exponent; junt amb estimadors geoestadístics: krigeado ordinari, krigeado simple, krigeado amb una deriva externa, cokrigeado simple, cokrigeado ordinari, cokrigeado colocalizado i cokrigeado colocalizado davall un model de Markov.

Finalment, per a possibilitar l’interacció d’un usuari de nivell intermedi amb els algoritmes implementats, es mostra una interfície gràfica d’usuari construïda per a gestionar i administrar els processos computacionals relacionats amb l’execució de les diferents anàlisis integrats a la llibreria, que permet a l’usuari realitzar l’ingrés de dades i l’extracció, visualització i anàlisi dels resultats. El programari (librerías i interfaz) desenrotllat es denomina ESTELLUS i està escrit completament en llenguatge de programació orientat a objectes Java per a possibilitar la implementació de la conceptualització proposada. A més de l’anterior, Java fa possible l’ús d’una gran quantitat de llibreries d’accés lliure existents, algunes de les quals van ser adaptades i utilitzades per a propòsits de visualització de resultats i anàlisi de dades. ESTELLUS pot ser usat efectivament en estudis de diferent tipus: modelació hidrogeològica, anàlisis hidrològiques, estudis ambientals, anàlisis meteorològiques, estudis de la ciència del sòl, anàlisi de risc, prospecció minera, etc. Per mitjà de l’aplicació de ESTELLUS per a la resolució d’un problema pràctic senzill en hidrogeologia es demostra la seua versatilitat per a l’anàlisi i visualització d’informació geogràfica i les seues àmplies possibilitats d’expansió futura en el concernent amb l’estimació i simulació estocàstica espacial.

Índice general

Agradecimientos	VII
Resumen	IX
Abstract	XI
Resum	XIII
1. Introducción	1
1.1. Motivación de este trabajo	1
1.2. Propósitos y alcances	3
1.3. Estructura del documento	4
2. Estado del arte	7
2.1. Introducción	7
2.2. Enfoques de la estimación espacial	7
2.2.1. Métodos deterministas	7
2.2.2. Métodos estocásticos	8
2.3. Características generales y evolución de las librerías computacionales geoes- tadísticas	9
2.3.1. Características y evolución del software para el cálculo de variogra- mas experimentales y ajuste de modelos	12
2.4. Comentarios finales	14
3. Generalidades y diseño de la librería	15
3.1. Introducción	15
3.2. Una visión general de ESTELLUS	15
3.2.1. ¿Porqué Java?	17
3.3. Librerías externas integradas para el desarrollo de ESTELLUS	17
3.3.1. HydroSIG	18
3.3.2. Jama	18
3.3.3. JFREECHART	19
3.3.4. VISAD	19

3.3.5. Librería matemática y científica del Dr. Flanagan	19
3.4. Vista general de la Interfaz de Programación de Aplicaciones (API) de ESTELLUS	20
3.5. Comentarios finales	27
4. Manejo de información espacial	29
4.1. Introducción	29
4.2. La referenciación espacial de las entidades geográficas	30
4.3. La manipulación de información puntual	32
4.3.1. La clase <code>Observacion</code>	33
4.3.2. La clase <code>ListaObservaciones</code>	33
4.4. La manipulación de información exhaustivamente conocida	34
4.4.1. Las clases <code>DataRaster</code> y <code>DataRaster3D</code>	35
4.4.2. Las clases <code>MetaRaster</code> y <code>MetaRaster3D</code>	37
4.5. Algoritmos para el análisis descriptivo de los datos	41
4.5.1. Análisis univariado	41
4.5.2. Análisis bivariado	41
4.6. Codificación indicadora de la información	42
4.6.1. La clase <code>VariableIndicadora</code>	42
4.6.2. La clase <code>VariableIndicadoraSuave</code>	42
4.7. Comentarios finales	43
5. Análisis estructural	45
5.1. Introducción	45
5.2. Estimación del variograma experimental directo	45
5.3. Estimación del variograma experimental cruzado	46
5.4. Parámetros para el cálculo de variogramas experimentales	46
5.5. Algoritmos para el cálculo de variogramas experimentales	47
5.5.1. Variogramas omnidireccionales	47
5.5.2. Variogramas direccionales	49
5.6. Variogramas superficiales	50
5.6.1. La clase <code>MapaVariograma</code>	50
5.7. Ajuste de modelos a los variogramas experimentales	52
5.7.1. Manejo computacional de los modelos de variograma	53
5.7.2. Ajuste mediante mínimos cuadrados ordinarios	54
5.7.3. Ajuste mediante mínimos cuadrados ponderados	54
5.7.4. El algoritmo de optimización <i>Nelder-Mead</i>	55
5.7.5. La clase <code>AjustarMinimosCuadradosNelderMead</code>	56
5.8. Modelos anisotrópicos de variograma	57
5.8.1. La clase <code>EvaluarVariograma</code>	59
5.8.2. La clase <code>EvaluarAnisotropia</code>	60
5.8.3. La clase <code>SuperposicionAnisotropia</code>	62

5.9. Comentarios finales	62
6. Estimación espacial	63
6.1. Introducción	63
6.2. Algoritmos de estimación determinista	63
6.2.1. Método del vecino cercano	63
6.2.2. Método del inverso de la distancia elevado a un exponente	64
6.2.3. Los objetos computacionales para la estimación determinista	64
6.3. Algoritmos de estimación estocástica de una variable	65
6.3.1. Kriging simple	66
6.3.2. Kriging ordinario	67
6.3.3. Los objetos computacionales para la estimación por krigeado simple y ordinario	67
6.4. Algoritmos de estimación que incorporan una tendencia	69
6.4.1. Krigeado con deriva externa	69
6.4.2. Los objetos computacionales para la estimación por krigeado con una deriva externa	70
6.5. Algoritmos de estimación que incorporan información externa	71
6.5.1. Cokrigeado simple	72
6.5.2. Cokrigeado ordinario	73
6.5.3. Cokrigeado ordinario estandarizado	74
6.5.4. Los objetos computacionales para la estimación por cokrigeado	74
6.6. Algoritmos de estimación que incorporan información externa exhaustivamente conocida	76
6.6.1. Cokrigeado simple colocalizado	77
6.6.2. Cokrigeado colocalizado bajo un modelo de Markov	77
6.6.3. Los objetos computacionales para la estimación mediante cokrigeado colocalizado bajo un modelo de Markov	78
6.7. Algoritmos de búsqueda	80
6.7.1. Búsqueda por bloques	81
6.7.2. Búsqueda por áreas vecinas	81
6.8. Validación cruzada	83
6.8.1. Las clases para efectuar la validación cruzada	83
7. Manual del usuario	85
7.1. Introducción	85
7.2. La interfaz principal	85
7.2.1. Menú de opciones	87
7.2.2. Lista de entidades geográficas	88
7.2.3. Funciones de manipulación de los campos raster	89
7.2.4. Panel de visualización de los campos raster	90

7.2.5.	Árbol de selección de algoritmos de estimación, panel de algoritmo y panel de ejecución	91
7.3.	Funciones de manipulación, análisis y visualización de campos raster	92
7.3.1.	Guardar imagen de campo	93
7.3.2.	Guardar campo raster seleccionado	93
7.3.3.	Exportar archivo raster bidimensional actual a formato <code>ascii</code> raster de ESRI	93
7.3.4.	Funciones generales de visualización y manipulación para campos raster	93
7.4.	Opciones de Archivo	99
7.4.1.	Abrir un archivo	100
7.4.2.	Entrada manual de información puntual	101
7.4.3.	Importar archivos puntuales	103
7.4.4.	Importar campos raster	104
7.4.5.	Importar desde archivos <code>ascii</code> raster de ESRI	106
7.4.6.	Exportar raster bidimensional a formato <code>ascii</code> de ESRI	107
7.5.	Herramientas para el análisis descriptivo de los datos	107
7.5.1.	Análisis estadístico univariado	107
7.5.2.	Análisis estadístico bivariado	109
7.6.	Herramientas para el cálculo de variogramas experimentales	111
7.6.1.	Cálculo de variogramas directos	112
7.6.2.	Cálculo de variogramas cruzados	116
7.7.	Herramientas para el ajuste de modelos de variogramas	119
7.7.1.	Ajuste de modelos a variogramas experimentales	119
7.7.2.	Creación de modelos anisotrópicos	121
7.8.	Herramientas para la estimación	124
7.8.1.	Acceso a los diferentes algoritmos de estimación	125
7.8.2.	Interfaz para el kriging ordinario y simple	125
7.8.3.	Interfaz para la estimación mediante los métodos deterministas	128
7.8.4.	Interfaz para el kriging con una deriva externa	128
7.8.5.	Interfaz para el cokriging ordinario y ordinario estandarizado	129
7.8.6.	Interfaz para la ejecución del cokriging colocalizado bajo un modelo de Markov	130
7.8.7.	Ejecución de la técnica de validación cruzada y visualización de sus resultados	132
8.	Ejemplo de aplicación	151
8.1.	Introducción	151
8.2.	La base de datos Berea	151
8.3.	Análisis descriptivo de la información	154
8.4.	Análisis estructural	158
8.5.	Estimación	163

Índice general XIX

8.5.1. Kriging ordinario y krigeado simple	163
8.5.2. Kriging con una deriva externa	163
8.5.3. Cokrigeado ordinario estandarizado	163
8.5.4. Cokrigeado colocalizado bajo un modelo de Markov	164
8.6. Análisis de las estimaciones por validación cruzada	165
8.7. Comentarios finales	168
9. Conclusiones	173
9.1. Líneas futuras de trabajo e investigación	175
Referencias	177

Índice de figuras

4.1.	Esquema conceptual de las funciones y propiedades de los objetos pertenecientes a la clase Georeferenciacion	31
4.2.	Esquema conceptual empleado para la representación del dominio espacial de un campo raster bidimensional, implementado en la clase MetaRaster	36
4.3.	Esquema conceptual empleado para la representación del dominios espacial de un campo raster tridimensional, implementado en la clase MetaRaster3D	38
5.1.	Esquema de algunos de los parámetros relativos al cálculo de variogramas experimentales para datos irregularmente distribuidos en el espacio.	47
5.2.	Esquema de los ángulos que definen la orientación del eje direccional del variograma	50
5.3.	Esquema de la primer rotación de los ejes coordenados para lograr que éstos coincidan con los ejes de anisotropía. Se define el ángulo α , denominado acimut, para rotar los ejes X y Y alrededor del eje Z . De esta forma se obtiene un nuevo sistema coordenado (X', Y', Z)	59
5.4.	Esquema de la segunda rotación de los ejes coordenados para lograr que éstos coincidan con los ejes de anisotropía. Se define el ángulo β , denominado dirección, para rotar los ejes X' y Z alrededor del eje Y' . De esta forma se obtiene un nuevo sistema coordenado (X'', Y', Z')	60
5.5.	Esquema de la tercera rotación de los ejes coordenados para lograr que éstos coincidan con los ejes de anisotropía. Se define el ángulo θ , denominado inmersión, para rotar los ejes Y' y Z' alrededor del eje X'' . De esta forma se obtiene un nuevo sistema coordenado (X'', Y'', Z'')	61
6.1.	Esquema de las estrategias de búsqueda implementadas en ESTELLUS. Búsqueda por bloques (izquierda) y por áreas vecinas (derecha).	80
7.1.	Interfaz principal de ESTELLUS. Visualización de un variograma superficial.	86
7.2.	Interfaz principal de ESTELLUS. Visualización de un campo raster en dos dimensiones.	86
7.3.	Interfaz principal de ESTELLUS. Visualización de un campo raster en tres dimensiones.	87

7.4. Barra de funciones disponibles para la manipulación de campos raster bidimensionales.	92
7.5. Barra de funciones disponibles para la manipulación de campos raster tridimensionales.	92
7.6. Imágenes almacenadas de campos raster bidimensionales y tridimensionales utilizando la herramienta guardar imagen.	94
7.7. Opciones de selección de paletas de colores predefinidas	96
7.8. Visualización tridimensional y de isolíneas de campos bidimensionales.	97
7.9. Interfaz para la estimación del histograma de un campo raster bidimensional.	98
7.10. Interfaz para la personalización de la paleta de colores de un campo raster.	100
7.11. Cuadro de diálogo para seleccionar el archivo a abrir.	101
7.12. Cuadro de diálogo para ingresar la información puntual y de campos raster.	102
7.13. Cuadro de dialogo para importar información puntual desde archivos en formato GEO-EAS.	104
7.14. Cuadro de diálogo para importar información raster desde archivos en formato GEO-EAS.	134
7.15. Interfaz implementada para importar archivos raster en formato ascii de ESRI.	135
7.16. Interfaz implementada para exportar archivos raster en formato nativo de ESTELLUS a formato ascii raster de ESRI.	136
7.17. Interfaz implementada que realiza el análisis estadístico descriptivo univariado en ESTELLUS.	137
7.18. Gráficos y tablas resultados de realizar un análisis descriptivo univariado para una en ESTELLUS.	138
7.19. Interfaz implementada que realiza el análisis estadístico descriptivo bivariado en ESTELLUS.	139
7.20. Gráficos y tablas resultados de realizar un análisis descriptivo bivariado para una en ESTELLUS	140
7.21. Interfaz diseñada en ESTELLUS para el cálculo de variogramas directos.	141
7.22. Interfaz diseñada en ESTELLUS para el cálculo de los variogramas cruzados.	142
7.23. Interfaz diseñada en ESTELLUS para el ajuste de modelos a un variograma experimental.	143
7.24. Interfaz diseñada en ESTELLUS para la creación de superposición lineal de modelos de variograma anisotrópicos.	144
7.25. Vista general del panel diseñado para la ejecución de los algoritmos de krigado ordinario y simple.	145
7.26. Panel diseñado en ESTELLUS para la entrada de los modelos de variograma, la información del dominio espacial del campo raster de salida y los parámetros del algoritmo de búsqueda para <i>KO</i> y <i>KS</i>	146
7.27. Vista general de los panel contenedores diseñados para la ejecución del algoritmos de estimación krigado con deriva externa <i>KDE</i>	147

7.28. Interfaces para la entrada de los modelos de continuidad espacial, dominio de salida y parámetros de búsqueda para el <i>CKO</i> y <i>CKOE</i>	148
7.29. Interfaz para la ejecución de los algoritmos de <i>CKO</i> y <i>CKOE</i>	149
7.30. Interfaz para la ejecución del algoritmos de <i>CKCM</i>	150
8.1. Campos de la permeabilidad (arriba) y de la porosidad (abajo) de la base de datos Berea.	152
8.2. Muestreo aleatorio estratificado de la permeabilidad (arriba) y valores de la porosidad correspondientes (abajo).	153
8.3. Resultados del análisis descriptivo univariado para las muestras de 64 datos de permeabilidad y porosidad.	155
8.4. Diagrama de dispersión para las 64 parejas de datos de permeabilidad y porosidad de Berea.	157
8.5. Variogramas direccionales de la permeabilidad (arriba), la porosidad (abajo) y cruzado (centro) para diferentes direcciones.	159
8.6. Variogramas superficiales de la permeabilidad (arriba), la porosidad (abajo) y cruzado (centro).	160
8.7. Modelo de corregeionalización para la permeabilidad (arriba izquierda), cruzado (arriba derecha) y la porosidad (abajo izquierda).	161
8.8. Variograma superficial de la permeabilidad (arriba), la porosidad (abajo) y cruzado (centro).	162
8.9. Campo de permeabilidad estimado mediante krigeado ordinario.	164
8.10. Campo de permeabilidad estimado mediante krigeado simple.	165
8.11. Campo de permeabilidad estimado mediante krigeado con deriva externa.	166
8.12. Campo de permeabilidad estimado mediante cokrigeado ordinario.	167
8.13. Campo de permeabilidad estimado mediante cokrigeado colocalizado bajo un modelo de Markov.	168
8.14. Histogramas para los valores de la permeabilidad estimados por validación cruzada.	169
8.15. Diagramas de dispersión de los valores de la permeabilidad estimados por validación cruzada vs. los valores registrados.	170

Índice de tablas

- 8.1. Estadísticos descriptivos univariados de la permeabilidad y la porosidad para las 64 localizaciones seleccionadas en el dominio de análisis 156
- 8.2. Tabla de cuantiles estimados para las muestras de permeabilidad y porosidad 156
- 8.3. Estadísticos derivados de la información obtenida por validación cruzada para las distintas técnicas de interpolación empleadas (Max: máximo, Min: mínimo, m: media, M: mediana, Var: varianza, σ : desviación estándar, CV: coeficiente de variación, CA: coeficiente de asimetría, K: coeficiente de kurtosis, ρ : Correlación de Pearson). 166

1

Introducción

1.1. Motivación de este trabajo

La solución de un problema relacionado con las ciencias de la tierra casi siempre implica llevar a cabo las tareas de exploración, descripción y predicción del comportamiento espacial de una o varias variables sobre el dominio de estudio. Por ejemplo, para cuantificar la cantidad disponible de agua en un lugar determinado el hidrólogo debe conocer la variabilidad de los procesos que la afectan, como la lluvia, la evapotranspiración y la infiltración, entre otros. De igual modo, si se desea determinar la cantidad de agua que puede ser extraída de un acuífero, se debe estimar el comportamiento espacial de las propiedades de los materiales que lo componen, entre ellas la conductividad hidráulica, la porosidad y la transmisividad.

Lograr “entender” la variabilidad espacial de los fenómenos naturales ha sido uno de los tópicos centrales de las ciencias de la tierra. Tarea en ningún modo sencilla puesto que, en general, los procesos estudiados son variables a diferentes escalas, tanto en el tiempo como en el espacio, y no suelen mostrar un comportamiento lineal. Además, las mediciones disponibles de las variables estudiadas suelen ser escasas. A pesar de lo anterior, si se analiza de forma sistemática la información disponible, es posible extraer patrones de comportamiento que constituyen, en sí mismos, una descripción de la variabilidad, y pueden ayudar al investigador a realizar una predicción de la variable analizada en regiones donde no se dispone de la misma. Bajo este marco conceptual, la Geoestadística ofrece técnicas para describir la estructura espacial de fenómenos naturales y construir/estimar representaciones válidas para conocer mejor los procesos y sistemas naturales (Isaaks y Srivastava, 1989; Gómez-Hernández, 1996).

Para ello la Geoestadística trata con información espacial conocida sobre localizaciones limitadas o, en algunos casos, con imágenes exhaustivas. En el primer caso la información puede asociarse con entidades geográficas puntuales, mientras que en el segundo caso la información puede conceptualizarse como un campo aleatorio espacial. Las entidades geográficas puntuales son objetos que pueden representarse de forma explícita mediante sus características espaciales y pueden tener asociados un conjunto de aspectos temáticos (Bosque-Sendra, 1997); mientras que, según Burrough y McDonnell (1998), un campo aleatorio es un modelo de datos que permite visualizar información geográfica que se caracteriza por ser continuamente variable en el espacio. De esta forma, un campo aleatorio puede considerarse como un modelo sencillo para representar la distribución espacial de variables. Ambos tipos de representación son compatibles con los conceptos de modelos de objetos (o entidades) espaciales utilizados ampliamente en los Sistemas de Información Geográfica (*SIG*), es decir, los formatos de datos vectorial y raster, respectivamente.

Parece clara entonces la relación entre los conceptos básicos de modelación espacial de objetos geográficos, la utilización de funciones de manipulación y análisis *SIG* y las técnicas geoestadísticas, la cual proporciona herramientas útiles en lo que se refiere a (Burrough, 2001):

- Geo-registro: el uso de un marco conceptual bajo las premisas de los *SIG* en los análisis geoestadísticos obliga a que las coordenadas de las localizaciones con información se encuentren referidas a un sistema de coordenadas pre-establecido o estandarizado. Lo anterior permite el uso de datos registrados en campo vía sistema de posicionamiento global, *GPS*.
- Análisis exploratorio de los datos: la Geoestadística proporciona herramientas para el análisis exploratorio espacial de los datos que pueden ser implementadas efectivamente utilizando conceptos *SIG*.
- Contexto espacial y uso de información externa: los *SIG* hacen posible la importación de datos que pueden ser usados como información secundaria en análisis, estimación y simulación geoestadística.
- Visualización: los *SIG* ofrecen herramientas para manipular los resultados de las estimaciones y permitir el despliegue visual de éstos con otras fuentes de información disponibles en el dominio de estudio.

Bajo este marco unificador, una herramienta computacional para un estudio geoestadístico debe proporcionar las herramientas de análisis y estimación que se encuentran frecuentemente en las aplicaciones de análisis de datos y en las librerías geoestadísticas y que, de acuerdo con ESRI (2001) serían:

1. Herramientas para el análisis exploratorio.
2. Herramientas para el análisis estructural.

3. Herramientas para la estimación.

Todo ello bajo un entorno computacional en el cual la localización de los objetos geográficos se represente mediante coordenadas pertenecientes a un sistema geográfico real.

Desde el punto de vista informático, la creación de una librería de algoritmos geoes-tadísticos que integre conceptos geográficos puede realizarse efectivamente utilizando la programación orientada a objetos. Este paradigma de programación considera a un programa como una colección de entidades autónomas llamados *objetos*. Cada objeto es responsable de realizar tareas específicas en la ejecución del programa. La interacción entre los objetos ocasiona que el programa avance en su ejecución, por lo tanto, resulta ser la simulación de un universo modelo (Budd, 1994). La incorporación de conceptos geográficos ocasiona que un objeto computacional virtual que se haya denominado como *espacial*, ineludiblemente presente un atributo relacionado con su localización. Dicho atributo hace posible la construcción de algoritmos computacionales para ejecución de funciones de superposición y análisis de tipo *SIG*.

La Geoestadística y los *SIG* proporcionan herramientas que pueden ser aplicadas efectivamente a cualquier rama de la ciencia relacionada con la modelación de fenómenos espaciales como las ciencias del suelo, la agricultura, la geología, la hidrología, la meteorología, la oceanografía, la geografía física, etc. Así, en el quehacer científico de dichas disciplinas es de inmenso valor disponer de una herramienta computacional, que permita a los investigadores realizar estimaciones utilizando algoritmos geoes-tadísticos, que sea flexible para adaptarse a dominios espaciales y que los resultados obtenidos puedan ser compatibles con los *SIG* comerciales.

1.2. Propósitos y alcances

El propósito principal de este trabajo es la exploración e implementación de algoritmos computacionales que le permitan al modelizador geoes-tadístico la inclusión explícita de la componente geográfica en los procedimientos de análisis de la continuidad espacial y la predicción de valores donde no se dispone de mediciones. La idea fundamental es conceptualizar objetos computacionales que representen las entradas y salidas de los algoritmos de estimación y asociar a éstos su localización de acuerdo con un sistema de coordenadas reales. Así los objetos de información manipulados, creados, estimados y analizados por las técnicas geoes-tadísticas son geográficos y pueden ser integrados a un análisis *SIG* mediante operaciones de importación y exportación de formatos en alguno de los software comerciales existentes.

El alcance del trabajo se enmarca dentro del desarrollo de las clases computacionales que manipulan la información geográfica, la implementación de algunas técnicas de estimación geoes-tadística y la creación de un entorno visual para que un usuario con experiencia media en geoes-tadística pueda realizar estimaciones de campos raster. En ningún momento este trabajo pretende realizar el desarrollo de un sistema de información geográfica, en el sentido amplio. Lo que se desea es llenar un vacío conceptual existente

en las librerías geoestadísticas actuales, como se presenta en el capítulo 2, mediante la construcción y adaptación de conceptos que integren la geografía y la geoestadística.

Para ello los temas tratados en este trabajo son:

- Una revisión bibliográfica de las librerías y programas geoestadísticos disponibles, con énfasis en las ventajas y carencias de sus algoritmos e interfaces de usuario.
- El planteamiento del diseño de la librería de clases programadas para realizar el análisis de la información y la estimación geoestadística, así como la interfaz gráfica de usuario.
- Una descripción de las librerías adaptadas y/o utilizadas en la construcción de los programas nativos que contienen los algoritmos que implementan las técnicas incluidas.
- Un análisis de las ventajas y desventajas del lenguaje de programación seleccionado, **Java**, para la implementación computacional de los algoritmos propuestos.
- Una descripción detallada de las propiedades que caracterizan a los objetos espaciales usados junto con sus implicaciones computacionales.
- Una discusión en detalle de las bases conceptuales de los algoritmos de estimación, los objetos diseñados y sus implicaciones computacionales. Esta cuestión permitirá que las personas interesadas puedan seguir trabajando en la construcción de algoritmos a partir de lo realizado en el presente trabajo.
- La presentación de un manual de usuario del software.
- Una aplicación práctica a un caso de estudio sencillo.

1.3. Estructura del documento

Para lograr lo anterior, este documento se divide en nueve capítulos. El primero es una introducción general. En el capítulo 2 se presenta un análisis acerca de las características generales de los softwares de estimación geoestadística existentes. En el capítulo 3 se presentan las consideraciones llevadas a cabo para el diseño de la librería computacional, las clases utilizadas y modificadas e integradas desde otras librerías y una visión general acerca de los componentes de la interfaz gráfica de ESTELLUS. En el capítulo 4 se discuten en detalle las clases implementadas para la representación de la información geográfica a integrarse en los procedimientos de estimación. En el capítulo 5 se muestran las bases conceptuales geoestadísticas usadas para la implementación de los algoritmos de análisis estructural y las clases programadas para dicho fin. En el capítulo 6 se detallan los algoritmos de estimación implementados y las clases programadas para su ejecución. En el capítulo 7 se presenta un manual de usuario para la interfaz gráfica de usuario

Capítulo 1. Introducción

5

desarrollada. En el capítulo 8 se presenta una breve aplicación de todas las utilidades del software. Finalmente, en el capítulo 9 se presentan las conclusiones del trabajo, algunas recomendaciones y las futuras líneas de investigación y desarrollo a seguir.

2

Estado del arte

2.1. Introducción

En este capítulo se presenta una revisión del estado del arte en lo referente a las más extendidas herramientas computacionales existentes para la interpolación espacial de variables. Inicialmente se presentan algunas ideas breves acerca de las bases conceptuales relacionadas con los algoritmos de estimación. Seguidamente, se comenta de manera resumida las características de las librerías de algoritmos y de las aplicaciones geoestadísticas más ampliamente usadas en la actualidad para la realización de estimaciones espaciales. Finalmente, se discuten las herramientas existentes para el análisis de los datos a usar en la estimación y construcción de modelos de variación espacial para los atributos que se desea estimar.

2.2. Enfoques de la estimación espacial

Los métodos de interpolación espacial utilizan información discreta de una variable de interés para obtener superficies continuas. De manera general, puede decirse que estos se clasifican en dos categorías: los métodos deterministas y los métodos estocásticos (Meijernik et al., 1994).

2.2.1. Métodos deterministas

Diferentes autores (Dubrule, 1983; Isaaks y Srivastava, 1989; Meijernik et al., 1994; Burrough y McDonnell, 1998; ESRI, 2001) presentan una revisión completa de los diver-

Los métodos deterministas clásicos para generar superficies continuas de variables espaciales a partir de valores medidos puntualmente, como por ejemplo el vecino cercano, el método lineal, el método del inverso de la distancia elevada a un exponente, los splines, la interpolación cúbica y bicúbica, etc. A pesar de que dichos métodos no pueden cuantificar la incertidumbre de interpolación, muchos autores manifiestan que éstos tienen la gran ventaja de ser algoritmos muy sencillos que pueden ser fácilmente implementados computacionalmente en librerías asociadas a sistemas de información geográfica.

2.2.2. Métodos estocásticos

El desarrollo de la Geoestadística en la década de los sesenta surgió de la necesidad de generar una metodología para evaluar y estimar la disponibilidad de reservas en los depósitos mineros. Posteriormente, debido a su facilidad para operar, analizar y extraer patrones de comportamiento de variables distribuidas en el espacio, fue adoptada como herramienta para el análisis de la variabilidad espacial de los procesos físicos en diversos campos de las ciencias de la tierra, como por ejemplo la oceanografía, la hidrogeología, la hidrología superficial, la meteorología, la ciencia del suelo, la investigación de yacimientos minerales, entre tantos (Journel y Huijbregts, 1978; Gómez-Hernández, 1996; Goovaerts, 1997).

Como uno de sus campos fundamentales de desarrollo, la Geoestadística ha generado diversos métodos de estimación, los cuales se conocen como modelos estocásticos o estadísticos de interpolación. Estos métodos consideran que la variable espacial que se desea interpolar se comporta en parte de forma aleatoria. Además, se considera que el conjunto de datos disponible es una muestra de un conjunto infinito posible de muestras, con la característica de que son eventos que cumplen con la misma función de distribución conjunta de probabilidad (Meijerik et al., 1994). Para fines de estimación, la Geoestadística considera que la variable que se desea estimar tiene intrínseca dos componentes, una componente determinista que varía uniformemente en el espacio y una componente aleatoria. La componente determinista permite hacer una afirmación acerca del rango probable de valores a los que pueden pertenecer los correspondientes a las zonas no conocidas, mientras que la componente aleatoria es el causante de que no puedan hacerse predicciones con total certidumbre (Gómez-Hernández, 1996; Burrough y McDonnell, 1998).

Cuando se emplea la Geoestadística como una herramienta para la interpolación de campos, se parte del ajuste de un modelo de variabilidad espacial de la variable representada (variograma, función de covarianza o correlograma). Basándose en dicho modelo de variabilidad espacial, se configura un estimador lineal no sesgado, el cual minimiza la varianza de estimación en un conjunto de estimaciones, es decir, se trata de algoritmos que utilizan la regresión espacial lineal de mínimos cuadrados. Estos métodos se conocen como krigado, en honor al trabajo de estimación espacial publicado por Daniel Krige en 1951 (Matheron, 1965, 1971; Journel y Huijbregts, 1978). Entonces, las técnicas geoestadísticas de interpolación o krigado, permiten efectuar evaluaciones del comportamiento espacial de variables regionalizadas, interpolación óptima de las mismas y estimaciones de la con-

fiabilidad de las estimaciones puesto que proporcionan un valor de la varianza del error de estimación, la cual puede interpretarse como una estimación de la incertidumbre de la estimación realizada (Journel y Huijbregts, 1978; Isaaks y Srivastava, 1989; Goovaerts, 1997; Burrough y McDonnell, 1998).

Igualmente, una de las mayores contribuciones de la Geoestadística en la modelación y la estimación espacial de campos que representan alguna variable ha sido, precisamente, el direccionamiento del problema de la estimación de variables espaciales hacia la integración de distintos tipos de datos, proponiendo para dicho fin una serie de algoritmos capaces de relacionar de forma sistemática la información obtenida a partir de distintas fuentes. En general, la información secundaria proviene de mediciones indirectas que pueden dar indicio de la distribución espacial de la variable principal que se desea modelar espacialmente y que se correlacionan con ésta (Cassiraga, 1999).

2.3. Características generales y evolución de las librerías computacionales geoestadísticas

Las primeras herramientas informáticas en Geoestadística fueron desarrolladas en la década de los ochenta para permitir la resolución de problemas de análisis y estimación espacial de variables y propiedades físicas. Entre ellas, GEO-EAS (EPA, 1991) y GSLIB (Deutsch y Journel, 1998) fueron las más ampliamente difundidas en el ámbito científico. En particular, GEO-EAS es una herramienta de interpolación geoestadística bidimensional y análisis estadístico descriptivo de la información. Su uso se hizo extenso debido a su carácter gratuito y la construcción de un archivo compilado ejecutable bajo DOS y distribuido por la *United States Environmental Protection Agency, EPA*. Su principal deficiencia era que su funcionalidad estaba limitada a dominios espaciales bidimensionales (Wingle et al., 1999).

Por su parte, GSLIB, desarrollada en el Departamento de Ingeniería del Petróleo de la Universidad de Stanford y publicada por primera vez en 1992, es una librería en todo el sentido de la palabra y proporciona al usuario un amplio conjunto de algoritmos para realizar análisis estructural, estimación y simulación geoestadística en dominios espaciales bidimensionales y tridimensionales. Algunos investigadores la consideran como el principal desarrollo informático en la historia de la Geoestadística (Remy, 2001). Debido a su diseño modular y su gran eficiencia ha sido empleada como herramienta de enseñanza e investigación en el desarrollo de nuevos algoritmos y metodologías (Deutsch y Journel, 1998; Remy, 2001). Su principal desventaja son las limitadas capacidades de visualización de resultados, puesto que sólo soporta la visualización mediante archivos Post-Script, además de la inexistencia de una interfaz gráfica *freeware* de usuario (Wingle et al., 1999).

A mediados de los años noventa surgieron muchas aplicaciones geoestadísticas en las cuales los análisis se desarrollaran de forma interactiva. Así, Carr y Mela (1998) traducen los programas FGAM y JBLOCK, escritos para el cálculo de variogramas y la estimación mediante krigeado y originalmente disponibles en FORTAN77, al lenguaje VisualBasic

de MICROSOFT con el fin de poder crear una interfaz gráfica de usuario que pueda ser desplegada en forma de ventanas en WINDOWS95.

Igualmente, desde una perspectiva más general, se desarrollaron paquetes como GSTAT (Pebesma y Wesseling, 1998) y UNCERT (Wingle et al., 1999), que ofrecen una amplia variedad de herramientas computacionales y de visualización, entre las cuales las más importantes son:

- Análisis estadístico descriptivo univariado y bivariado.
- Estimación de variogramas experimentales y ajuste interactivo de modelos.
- Interfaces para la ejecución de algoritmos de estimación y simulación.

A pesar de tener ciertos rasgos comunes, éstos programas fueron creados para diferentes propósitos. UNCERT se concibió como una herramienta para estimar o simular información que pueda importarse dentro de modelos de flujo subterráneo, como MODFLOW (McDonald y Harbaugh, 1984), o de transporte de masa, como MT3D (Zheng, 1990). Por otro lado, GSTAT fue concebido para ser utilizado como una herramienta geoestadística en estudios ambientales en los cuales la información se obtienen de bases de datos muy grandes y los resultados se almacenan, procesan e imprimen empleando SIG. Para lo anterior, los autores incorporan el soporte de muchos tipos de formatos raster compatibles con sistemas de información geográfica como ARCINFO e IDRISI (Pebesma y Wesseling, 1998). Ambas librería presentaban la limitación de que su ejecución se debía realizar bajo el sistema operativo UNIX, por lo cual los ejecutables disponibles en esa época no eran compatibles para su ejecución en WINDOWS, limitando su uso a la comunidad científica y académica.

Como puede verse, la evolución de las librerías geoestadísticas está íntimamente relacionada con la evolución de los lenguajes de programación disponibles en las diferentes épocas, e igualmente implica la adopción de diferentes paradigmas de programación. Por ejemplo, GSLIB, desarrollado empleando el lenguaje FORTRAN77 y posteriormente re-escrita en FORTRAN90 en una librería que adopta el paradigma de la programación secuencial, heredada desde la conceptualización misma del lenguaje FORTRAN. A pesar de que se ha comprobado la eficiencia computacional del lenguaje de programación (García-Merayo, 1999), su carencia de estructuración, en especial el uso de las sentencias de salto de línea, ocasiona que los programas sean de difícil comprensión, inclusive para su propio autor (Schildt, 2001). Aún así, la utilidad de una librería como GSLIB sigue vigente debido a que puede afrontar problemas altamente demandantes en operaciones matemáticas.

De la misma manera, C fue uno de los lenguajes más utilizados a principios de la década de los noventa para el desarrollo de librerías científicas. Como caso especial, en la Geoestadística se desarrollaron alguna herramientas ya referidas previamente, como lo son GSTAT y UNCERT escritas en ANSI-C. C adopta el paradigma de programación estructurada lo cual presenta ciertas ventajas:

- Permite a lo programadores escribir fácilmente programas de complejidad moderada.

- Posibilita obtener la máxima velocidad con un mínimo uso de memoria.
- Proporciona facilidades para realizar programas modulares y/o utilizar código o bibliotecas existentes.

pero también tiene importantes inconvenientes como:

- No se puede escribir código independiente de la máquina.
- Se genera una mayor dificultad en la programación y en la comprensión del código cuando se trabaja con proyectos de tamaño considerable con respecto a los lenguajes orientados a objetos.

Para salvar las dificultades anteriores pueden usarse lenguajes de alto nivel orientados a objetos como C++ o Java, los cuales se caracterizan por expresar los algoritmos de una manera adecuada a la capacidad cognitiva humana, en lugar de a la capacidad de ejecución de las máquinas. La primera librería geoestadística que incorpora el lenguaje de alto nivel orientado a objetos C++ para la escritura de sus algoritmos es GSTL (Remy, 2001). El propósito de su desarrollo fue crear una librería que pudiese ser usada tanto para desarrollos de investigación como para aplicaciones directas, fácilmente extensibles y que fuera una solución de compromiso entre eficiencia y entendimiento de código. La librería está compuesta por unas bibliotecas de fundación, descritas extensamente por Ying (2001), que contienen las clases que crean objetos comunes para los diferentes tipos de algoritmos de estimación y simulación, y por los programas geoestadísticos propiamente dichos, los cuales se describen en Remy (2001) y en Ying (2002). Su diseño se ha realizado utilizando programación genérica, de forma que la implementación de los algoritmos se realice con la menor cantidad de información necesaria (Remy, 2001).

En base a los programas contenidos en la librería GSTL, Remy (2004) creó un nuevo software geoestadístico, llamado SGEMS, con dos fines principales:

1. Proporcionar una interfaz gráfica de usuario para facilitar el uso de los algoritmos de GSTL.
2. Diseñar un software cuyas capacidades puedan crecer mediante la escritura de *plugins*.

Igualmente, el proyecto GSTAT no fue ajeno a las nuevas tendencias de conceptualización de programación en las ciencias informáticas y a la evolución de los lenguajes de programación. Por ello, hizo compatible el código original del paquete GSTAT del lenguaje C, considerado de nivel medio, con las capacidades proporcionadas por el lenguaje S, que puede considerarse como un lenguaje de alto nivel. Pebesma (2004) afirma que esto se decidió debido a que dicho lenguaje proporciona funciones gráficas avanzadas para el despliegue de gráficos y mapas, los cuales consideraba limitados en la versión original.

Otro lenguaje de programación de alto nivel que empieza a ser ampliamente utilizado para el desarrollo de aplicaciones espaciales es Java. Por ejemplo, Rühaak (2006) presenta una aplicación desarrollada para efectuar interpolaciones espaciales tridimensionales

con krigeado ordinario y una versión modificada del método del inverso cuadrático de la distancia. Markstrom et al. (2002) afirma que la principal ventaja de utilizar **Java** en la modelación de fenómenos espaciales es la posibilidad de distribuir los programas creados mediante aplicaciones en internet debido a que es un lenguaje multiplataforma.

El desarrollo de las herramientas geoestadísticas puede también analizarse desde otro punto de vista, un poco menos técnico en lo que al conocimiento de la ciencia informática se refiere. Con el auge del desarrollo de los *SIG*, éstos se han convertido en una de las herramientas más eficaces para la estimación de campos de variables físicas a partir de la aplicación de diferentes tipos de procedimientos de manipulación de información espacial, entre los cuales se encuentran la interpolación, la superposición algebraica, el remuestreo o *resampling*, la suavización de texturas y la reclasificación de imágenes de satélite y campos en formato raster y, obviamente, la Geoestadística (Álvarez, 2007). Bajo esta perspectiva de investigación se han desarrollado aplicaciones geoestadísticas orientadas a la integración de sus algoritmos de análisis de datos y estimación de mapas raster en *SIG* comerciales. ESRI (2001) desarrolló una extensión de este tipo integrada en ARCGIS, enfocada a la estimación de campos bidimensionales mediante krigeado y cokrigeado, análisis estructural y diagnóstico de los resultados obtenidos mediante validación cruzada. A pesar de ser una herramienta potente, su uso está limitado a dos dimensiones, además, el costo de una licencia para su uso comercial es muy elevado.

2.3.1. Características y evolución del software para el cálculo de variogramas experimentales y ajuste de modelos

La modelación de la continuidad espacial, conocida como variografía o análisis estructural, tiene una importancia especial en un estudio geoestadístico puesto que (EPA, 1991; Chu, 1993):

- Permite obtener un modelo matemático para la distribución espacial de la variable.
- Puede revelar la complejidad de patrones espaciales, los cuales pueden llegar a indicar que tipo de algoritmo geoestadístico debe utilizarse para la representación del fenómeno estudiado de cara a su modelación matemática.

Por lo anterior y debido a que el análisis estructural es la parte más difícil y tediosa en un análisis geoestadístico, especialmente en tres dimensiones, la implementación de una interfaz gráfica es una herramienta muy útil para el analista (Chu, 1993).

Actualmente existe un amplio número de herramientas informáticas que hacen posible realizar el análisis estructural. En ellas, en general, éste procedimiento se divide en dos pasos:

1. El cálculo de los variogramas experimentales a partir de la información disponible, usualmente empleando el estimador de media (Cressie, 1985).
2. El ajuste de un modelo a los variogramas experimentales de forma manual o automática.

A pesar de lo anterior, dichas librerías presentan muchas diferencias en lo referido a las utilidades gráficas y el tipo de algoritmos que proporcionan al usuario.

GSLIB (Deutsch y Journel, 1998), contiene herramientas potentes para efectuar las estimaciones de variogramas experimentales y variogramas superficiales para variables en dominios espaciales tridimensionales, pero la entrada de datos y salida de los resultados se realiza en base a archivos, lo cual puede llegar a ser bastante confuso para el usuario. Adicionalmente, para obtener gráficos de los resultados, los archivos deben someterse a programas de post-procesamiento. Una solución al anterior problema fue propuesta por Chu (1993), quien construye el programa XGAM, una interfaz de usuario que implementa ciertos programas diseñados especialmente para recibir los parámetros de ejecución, manualmente o desde archivo, y posibilitar al ejecución de los algoritmos contenidos en GSLIB. Los resultados se despliegan de forma visual e interactiva para el usuario y pueden almacenarse en disco. El uso actual de XGAM es poco extendido puesto que fue diseñado para UNIX.

En general, como se ha comentado, las aplicaciones geoestadísticas proporcionan las herramientas algorítmicas y visuales para que el usuario estime el variograma experimental y obtenga un modelo ajustado mediante mínimos cuadrados ordinarios o ponderados. Así, usando herramientas como GSTAT (Pebesma y Wesseling, 1998), SGEMS (Remy, 2004) y ARCGIS un usuario con conocimientos básicos de geoestadística puede obtener de manera relativamente sencilla modelos de (cor)regionalización espacial para la(s) variable(s) de análisis, puesto que proporcionan interfaces de usuario que facilitan dicha labor. Por ejemplo, GSTAT esta enfocado al ajuste automático del modelo de variograma, mientras que SGEMS proporciona una herramienta flexible que le permite al usuario ajustar de forma manual el modelo de variograma.

En la literatura técnica, también pueden encontrarse herramientas computacionales enfocadas exclusivamente a la modelación del variograma y a su ajuste automático. A pesar de que el ajuste manual del modelo de variación espacial permite la inclusión del conocimiento experto del modelador, el ajuste automático se considera útil por diferentes razones (Pardo-Igúsquiza, 1999):

- Puede usarse como una herramienta para obtener una primera aproximación.
- Es muy útil cuando desean ajustarse muchos variogramas.
- Es un análisis objetivo puesto que diferentes analistas obtienen el mismo resultado.
- Permite cuantificar la incertidumbre del ajuste.

Pardo-Igúsquiza (1999) y Faulkner (2002) desarrollaron dos herramientas enfocadas al ajuste automático del variograma experimental. VARFIT (Pardo-Igúsquiza, 1999) es un programa escrito en lenguaje FORTRAN77 que ajusta automáticamente modelos a los variogramas experimentales mediante mínimos cuadrados ponderados empleando el algoritmo de optimización *Simplex* propuesto por Nelder y Mead (1965) y diferentes expresiones para los valores de las ponderaciones, como por ejemplo las propuestas por Cressie (1985) y

Zhang et al. (1995). Su principal falencia es que no proporciona una interfaz de interacción programa-usuario. Por otro lado, Faulkner (2002) propone una librería **Java** para el ajuste de modelos de variograma cuya finalidad principal es eliminar la subjetividad del analista de este tipo de análisis. Para ello implementa la metodología conceptual propuesta por Müller (1999), la cual se basa en el ajuste a partir de la nube de variograma de forma iterativa. Las operaciones matriciales necesarias en el procedimiento de ajuste se realizan utilizando las clases de la librería **Jama** y los resultados se despliegan de forma visual.

2.4. Comentarios finales

Existen diferentes tipos de aplicaciones geoestadísticas que pueden ser usadas efectivamente en las estimación de variables. Su utilización depende de las características del problema que se desea resolver, por ejemplo:

- **GSLIB** requiere que el usuario posea conocimientos de programación para compilar y ejecutar el código.
- **UNCERT** y **GSTAT** no son librerías de uso extensivo puesto que requieren el uso de sistemas operativos como **UNIX**, **LINUX** o **DOS**.
- **SGEMS** es un paquete potente, pero su versión actual presenta inconsistencias en lo que respecta a algunos programas y otros no se encuentran habilitados. Por otro lado, no incorpora la concepción geográfica del problema de la estimación espacial.
- La adquisición de una licencia de **ARCGIS** es bastante costoso. Su módulo geoes-tadístico sólo permite el tratamiento de dominios espaciales bidimensionales, por lo cual su uso en disciplinas que requieren el tratamiento de dominios tridimensionales, como la Hidrogeología es muy limitado.

En este trabajo se pretende conceptualizar una librería de clases que reúna las fortalezas de los softwares revisados aunque por ahora su alcance está limitado al procedimiento de interpolación o estimación espacial.

3

Generalidades y diseño de la librería

3.1. Introducción

En esta sección se presenta una visión general de la aplicación desarrollada en este trabajo, la cual se ha denominado ESTELLUS, de la siguiente manera: ES es el acrónimo de ESTimación ESpacial, TELLUS es el nombre de la diosa romana de la tierra, cuyo análogo en la mitología griega es Gea.

3.2. Una visión general de ESTELLUS

ESTELLUS es un software para la estimación espacial geoestadística construido completamente en lenguaje de programación Java. La librería de clases que lo compone está dividida en dos partes, una que contienen las clases que ejecutan los algoritmos computacionales y otra que contiene las interfaces gráficas que permiten al usuario introducir la información necesaria para la ejecución de los algoritmos. En los paquetes creados se pueden localizar como las carpetas `*.objetos.*` y `*.widgets.*` respectivamente. Por ejemplo, la clase que contiene el algoritmo para realizar un krigado ordinario bidimensional, a la cual se ha denominado como `KrigingOrd2D` se encuentra dentro del paquete `JGeoST.estimadores.univaridos.objetos`, mientras que la interfaz gráfica de usuario que lo ejecuta se encuentra en el paquete `JGeoST.estimadores.univaridos.widgets`.

La creación del software surge de la necesidad de disponer de una librería computacional junto con una herramienta gráfica que permitan realizar estimaciones bidimension-

ales y tridimensionales de campos espaciales en formato raster bajo un contexto georeferenciado, cuyos resultados puedan ser integrados efectivamente dentro de un *SIG*. El software se ha desarrollado de forma que sea de fácil acceso y comprensión tanto para usuarios expertos en geoestadística como para analistas de conocimiento intermedio. Como se ha discutido en el capítulo 2, una librería de este tipo no se encuentra disponible en la actualidad.

Las características principales de ESTELLUS son las siguientes:

- Está escrita completamente en lenguaje de programación **Java**, por lo cual es multiplataforma.
- Los datos de entrada a los algoritmos se tratan como objetos geográficos para los cuales existen metadatos asociados y pueden relacionarse a un dominio espacial de acuerdo con sistemas de coordenadas estándar reales. Para las entidades puntuales se utiliza una concepción vectorial sencilla con un atributo asociado. Para la información exhaustiva se adopta una representación geográfica mediante campos en formato raster.
- La información de salida correspondiente a los algoritmos de estimación son campos en formato raster y muestras puntuales de la variable analizada, las cuales han sido estimadas mediante la ejecución de un procedimiento de validación cruzada.
- Se incluyen algoritmos para el análisis estadístico descriptivo univariado y bivariado.
- Se incluyen algoritmos para el análisis estructural, esto es, cálculo de variogramas experimentales, directos y cruzados, omnidireccionales y direccionales y para el ajuste automático de modelos lineales de (co)regionalización.
- Se incluye la opción de cálculo de variogramas superficiales.
- Con el propósito de que el software pueda ser compatible con los *SIG* comerciales y las librerías geoestadísticas más utilizadas, se han implementado herramientas para la importación de datos y exportación de resultados a diferentes formatos estándar.
- Para posibilitar el análisis y la manipulación de los resultados obtenidos, se implementan algunas funciones sencillas de manipulación y análisis de campos raster bidimensionales y tridimensionales.
- Los algoritmos están integrados con una interfaz gráfica de usuario principal, la cual realiza las gestiones de ejecución de los programas, manejo de los objetos geográficos y despliegue de la información y resultados.

3.2.1. ¿Porqué Java?

El lenguaje de programación **Java**, lanzado al mercado en el año 1996, ha recibido enorme atención por parte de la comunidad desarrolladora de software y ha sido considerado como el lenguaje más ampliamente difundido en la historia. Es bien conocido que fue diseñado para su uso en internet puesto que ofrece clases que pueden importar información a través de redes y que puede embeberse fácilmente en páginas **HTML** en forma de **applets**. Junto con lo anterior, **Java** manifiesta grandes capacidades y herramientas para implementar algoritmos computacionales e interfaces gráficas de usuario, junto con el poder de estar diseñado especialmente para realizar programación orientada a objetos (Faulkner, 2002). La principal ventaja de **Java** es que es un lenguaje diseñado para el uso en cualquier tipo de plataformas, desde hornos microondas hasta **CPU's**, por lo cual es altamente portable siendo utilizable en cualquier tipo de procesador. Además, la sintaxis de programación es bastantes similar a la de **C++**, con lo que los programadores familiarizados con este lenguaje pueden adoptar **Java** de manera sencilla. Es un lenguaje muy robusto puesto que comprueba la sintaxis de los códigos en los procesos de compilación y ejecución, lo que hace prácticamente imposible que se tengan errores difíciles de localizar (Schildt, 2001). Así, en casos en los cuales utilizando **C++** se puede crear un código más eficiente, en **Java** se pueden utilizar técnicas mucho más avanzadas que simplifican la creación de software de forma que se disponga de más tiempo para una oportuna optimización.

En el caso de software desarrollado para la modelación relacionada con las ciencias de la tierra, Markstrom et al. (2002) muestran las ventajas que tiene **Java**. Un ejemplo lo podemos ver con lo ocurrido con el US Geological Survey's (USGS) Modular Modeling System (Leavesley, 1997). Esta herramienta fue escrita originalmente en **FORTRAN** y luego pasada a **Java** lo que permitió su mantenimiento para los diferentes sistemas operativos existentes (**WINDOWS**, **SOLARIS**, **LINUX**, **HP-UX** y **AIX**) con el consiguiente ahorro de tiempo y recursos humanos. Además se creó un **applet** para habilitar su uso a través de internet.

En lo que a este trabajo concierne, se pretende que la librería diseñada sea un proyecto de investigación que involucre la exploración de diferentes técnicas de análisis, estimación y simulación geoestadística. Para ello será indispensable la participación de investigadores y programadores, que deberán encontrarse con un código sencillo, de fácil comprensión y portable, características inequívocas de **Java**, de manera que el software pueda ser fácilmente asimilado, modificado y ampliado.

3.3. Librerías externas integradas para el desarrollo de ESTELLUS

Existen en la literatura científica gran cantidad de esfuerzos de programación de uso libre que pueden ser empleados, modificados y adaptados para las necesidades particulares de los investigadores. A continuación se presenta una descripción de las librerías bajo

licencia GPL (*General Public License*, <http://www.gnu.org/licenses/gpl.html>) utilizadas para el desarrollo de ESTELLUS.

3.3.1. HidroSIG

HIDROSIG es un sistema de información geográfica con énfasis en la modelación hidrológica y la simulación de los recursos ambientales desarrollado en el marco del Postgrado en Aprovechamiento de Recursos Hidráulicos en la Universidad Nacional de Colombia Sede Medellín (Poveda et al., 2007 a,b). Las fortalezas de esta librería radican principalmente en que proporciona clases que manipulan objetos geográficos de manera efectiva, en especial campos aleatorios raster. Para el desarrollo de ESTELLUS se han adoptado y modificado algunas de las clases relacionadas con el manejo de información geográfica como son aquellas que tienen que ver con la manipulación de los campos en formato raster: `DataRaster` y `MetaRaster`, y la localización espacial: `Georeferenciacion`. Las modificaciones realizadas a dichas clases se describen extensamente en el capítulo 4.

3.3.2. Jama

En ESTELLUS se ha utilizado la librería matemática JAMA que es un paquete de álgebra lineal desarrollado por la empresa MATHWORKS. Está escrito en Java y proporciona clases para construir y operar sobre matrices de gran tamaño en formato de número `double`. Éstas han sido construidas de forma que los algoritmos implementados puedan enfrentar problemas complejos de estabilidad numérica cuando se utilizan para la solución de sistemas lineales altamente inestables. JAMA se ha concebido para servir como la librería estándar para el manejo de matrices para Java y es un código de uso bajo licencia pública de libre distribución (GPL) que puede descargarse de forma gratuita vía internet en la página web <http://math.nist.gov/javanumerics/jama/>. La librería forma parte del esfuerzo conjunto de muchos investigadores en el campo de la computación numérica mediante Java, denominado *Java Numerics*. Para más información el lector interesado puede consultar la página <http://math.nist.gov/javanumerics/>.

JAMA está formada por las clases: `Matrix`, `CholeskyDecomposition`, `LUDecomposition`, `QRDecomposition`, `EigenvalueDecomposition` y `SingularValueDecomposition`. ESTELLUS sólo emplea `Matrix`, que es la clase que dispone de algoritmos eficientes para invertir matrices de valor real y resolver sistemas de ecuaciones lineales, los cuales pueden consultarse en el manual de computación automática de Wilkinson y Reinsch (1972), y que han sido ampliamente utilizados en paquetes informáticos especializados en matemáticas, como MATLAB y LINPACK. Los métodos de la clase `Matrix` utilizados por ESTELLUS son `inverse` y `solve`. El objeto de la clase se construye ingresando como parámetro una matriz sobre la cual operan los métodos de clase.

3.3.3. JFreeChart

JFREECHART es una librería creada por la organización JFree (2005), que permite realizar el despliegue de gran cantidad de tipos de gráficos: histogramas, diagramas de dispersión, gráficos de barras bidimensionales y tridimensionales, gráficos circulares, gráficos de funciones y series temporales, diagramas de Gantt, entre muchos. Tiene la ventaja de ser bastante flexible para presentar los gráficos y de integrar opciones interactivas a éstos, como lo son el zoom, cambios de escalas de los ejes y la posibilidad de almacenar imágenes en disco. Lo anterior permite que durante la ejecución de la aplicación que actualmente use JFREECHART, el usuario pueda realizar la exploración interactiva de los resultados sobre las gráficas que han sido desplegadas. ESTELLUS usa las clases relacionadas con la construcción de histogramas, diagramas de dispersión, funciones de probabilidad acumulada experimental y, en general, gráficos de los resultados de los análisis descriptivos de la información. El lector interesado en conocer más a fondo las utilidades de esta librería puede remitirse a la página web del proyecto: www.jfree.org/jfreechart/.

3.3.4. VisAD

VISAD es una librería de algoritmos computacionales enfocada en la manipulación, despliegue bidimensional y tridimensional y análisis de datos científicos, en la que se definen constructores de clase y métodos para describirlos y desplegarlos vía metadatos, los cuales documentan la relación de éstos con el mundo que los rodea. La combinación de esquemas de datos y el despliegue de mapas permite obtener la visualización deseada (Hibbard, 1995, 1998, 2002). Las diferentes utilidades de VISAD han sido empleadas de forma efectiva por diferentes aplicaciones como por ejemplo HIDROSIG (Poveda et al., 2007 a,b), IDV (Merteens et al., 2004, 2005; Merteens y Wier, 2006; Wier, 2008) y CUENCAS (Mantilla y Gupta, 2005) para el despliegue de los mapas raster de diferentes tipos de información. Particularmente, en ESTELLUS las clases de VISAD se han utilizado para la construcción de la visualización de campos raster bidimensionales y tridimensionales.

3.3.5. Librería matemática y científica del Dr. Flanagan

Michael Flanagan (2008), profesor de UCL London’s Global University, lleva varios años desarrollando una librería científica usando Java en colaboración con sus alumnos. La librería es muy completa, cuenta con funcionalidad para realizar operaciones con matrices, estadísticas, generación de números aleatorios, propagación de errores, regresión, optimización, transformadas de Fourier, integrales, interpolaciones, buscar raíces, operaciones con números complejos, etc. Sus puntos débiles son que ha sido desarrollada por científicos, y no por ingenieros de software. La información acerca de los algoritmos, las clases y los métodos implementados en dicha librería se pueden consultar en <http://www.ee.ucl.ac.uk/mflanaga/java/>. Dentro de la librería existe una clase denominada **Stat**, en la cual están contenidos los diferentes algoritmos para la estimación de los estadísticos de una muestra de datos. Los estadísticos descriptivos usados en ESTELLUS son:

máximo, mínimo, rango, media (geométrica, armónica y aritmética), varianza, desviación estándar, coeficiente de asimetría, coeficiente de kurtosis y coeficiente de variación. Igualmente, se adaptó la clase `Regression` para lograr el ajuste automático de los modelos permitidos de variograma.

3.4. Vista general de la Interfaz de Programación de Aplicaciones (API) de ESTELLUS

A continuación se presenta de forma elemental la API de ESTELLUS. La biblioteca de clases está contenida dentro de los paquetes que se enumeran a continuación.

Paquete `JGeoST.descriptiva`

A Descripción: este paquete contiene las clases que realizan el análisis descriptivo de los datos.

B Lista de clases:

Clase `ProbabilidadExperimental`: calcula la probabilidad experimental de un valor muestral.

Clase `AnálisisUnivariado`: es un `JFrame` que contiene los métodos que realizan el cálculo de los estadísticos univariados y la información a desplegar en las gráficas correspondientes.

Clase `AnálisisBivariado`: es un `JFrame` que contiene los métodos que realizan el cálculo de los estadísticos bivariados y la información a desplegar en las gráficas correspondientes.

Clase `BibujarBivariado`: es un `JDialog` que despliega los resultados del análisis bivariado.

Clase `DibujarHistograma`: es un `JDialog` que despliega los resultados del análisis univariado.

Paquete `JGeoST.error`

A Descripción: este paquete contiene las clases que realizan la gestión de errores en la librería.

B Lista de clases:

Clase `JGeoSTException`: esta clase hereda de la clase `Exception` y permite realizar las gestiones de error en el interior de ESTELLUS.

Paquete `JGeoST.estimadores`

A Descripción: este paquete contiene las clases que efectúan las predicciones de un atributo empleando diferentes estimadores espaciales.

B Lista de clases:

- Clase KrigingOrd2D:** permite realizar la estimación de campos bidimensionales utilizando el algoritmo de krigeado ordinario.
- Clase KrigingOrd3D:** permite realizar la estimación de campos tridimensionales utilizando el algoritmo de krigeado ordinario.
- Clase KrigingSim2D:** permite realizar la estimación de campos bidimensionales utilizando el algoritmo de krigeado simple.
- Clase KrigingSim3D:** permite realizar la estimación de campos tridimensionales utilizando el algoritmo de krigeado simple.
- Clase KrigingExtDrf2D:** permite realizar la estimación de campos bidimensionales utilizando el algoritmo de krigeado con deriva externa.
- Clase KrigingExtDrf3D:** permite realizar la estimación de campos tridimensionales utilizando el algoritmo de krigeado con deriva externa.
- Clase CokrigingOrd2D:** permite realizar la estimación de campos bidimensionales utilizando el algoritmo de cokrigeado ordinario.
- Clase CokrigingOrd3D:** permite realizar la estimación de campos tridimensionales utilizando el algoritmo de cokrigeado ordinario.
- Clase CokrigingEst2D:** permite realizar la estimación de campos bidimensionales utilizando el algoritmo de cokrigeado estandarizado.
- Clase CokrigingEst3D:** permite realizar la estimación de campos tridimensionales utilizando el algoritmo de cokrigeado estandarizado.
- Clase CokrigingCol2D:** permite realizar la estimación de campos bidimensionales utilizando el algoritmo de cokrigeado colocalizado.
- Clase CokrigingCol3D:** permite realizar la estimación de campos tridimensionales utilizando el algoritmo de cokrigeado colocalizado.
- Clase CokrigingColMarkov2D:** permite realizar la estimación de campos bidimensionales utilizando el algoritmo de cokrigeado colocalizado markoviano.
- Clase CokrigingColMarkov3D:** permite realizar la estimación de campos tridimensionales utilizando el algoritmo de cokrigeado colocalizado markoviano.
- Clase VecinoCercano:** permite realizar la estimación de campos utilizando el algoritmo del vecino cercano.
- Clase InversoCuadratico:** permite realizar la estimación de campos utilizando el algoritmo de la inversa del cuadrado de la distancia.
- Clase ValidaciónCruzadaK0:** efectúa la validación cruzada para una estimación mediante krigeado ordinario.
- Clase ValidaciónCruzadaKS:** efectúa la validación cruzada para una estimación mediante krigeado simple.
- Clase ValidaciónCruzadaKDE:** efectúa la validación cruzada para una estimación mediante krigeado con deriva externa.

- Clase ValidaciónCruzadaCKO:** efectúa la validación cruzada para una estimación mediante cokrigado ordinario.
- Clase ValidaciónCruzadaCKE:** efectúa la validación cruzada para una estimación mediante cokrigado estandarizado.
- Clase ValidaciónCruzadaCKC:** efectúa la validación cruzada para una estimación mediante cokrigado colocalizado.
- Clase ValidaciónCruzadaCKCM:** efectúa la validación cruzada para una estimación mediante cokrigado colocalizado markoviano.
- Clase KrigingSearchEstrategy2D:** permite realizar la estrategia de búsqueda en la ejecución de un krigeado bidimensional (o de algún otro estimador univariado).
- Clase KrigingSearchEstrategy3D:** permite realizar la estrategia de búsqueda en la ejecución de un krigeado tridimensional (o de algún otro estimador univariado).
- Clase OrederingOperation2D:** realiza la búsqueda bidimensional en espiral como procedimiento de reclasificación previa de datos para optimizar la interpolación.
- Clase OrederingOperation3D:** realiza la búsqueda tridimensional en espiral como procedimiento de reclasificación previa de datos para optimizar la interpolación.
- Clase JFrameKO:** es un JFrame que permite la entrada de los parámetros necesarios para ejecutar la estimación mediante krigeado ordinario.
- Clase JFrameKS:** es un JFrame que permite la entrada de los parámetros necesarios para ejecutar la estimación mediante krigeado simple.
- Clase JFrameKDE:** es un JFrame que permite la entrada de los parámetros necesarios para ejecutar la estimación mediante krigeado con deriva externa.
- Clase JFrameCKO:** es un JFrame que permite la entrada de los parámetros necesarios para ejecutar la estimación mediante cokrigado ordinario.
- Clase JFrameCKE:** es un JFrame que permite la entrada de los parámetros necesarios para ejecutar la estimación mediante cokrigado estandarizado.
- Clase JFrameCKC:** es un JFrame que permite la entrada de los parámetros necesarios para ejecutar la estimación mediante cokrigado colocalizado.
- Clase JFrameCKCM:** es un JFrame que permite la entrada de los parámetros necesarios para ejecutar la estimación mediante cokrigado colocalizado markoviano.

C Lista de interfaces:

- Interfaz Kriging2D:** clase abstracta que permite la conceptualización de un estimador del tipo krigeado bidimensional.
- Interfaz Kriging3D:** clase abstracta que permite la conceptualización de un estimador del tipo krigeado tridimensional.

Paquete JGeoST.exportar

A Descripción: este paquete contiene las clases que efectúan las operaciones de exportación de formatos.

B Lista de clases:

Clase ASCII: permite la exportación de un campo raster en formato ESTELLUS a formato ascii raster de ESRI.

C Lista de interfaces:

Interfaz Exportar: es una clase abstracta que conceptualiza el procedimiento de exportación de un campo raster.

Paquete JGeoST.georef

A Descripción: este paquete contiene las clases que definen el sistema de georeferenciación de las entidades espaciales.

B Lista de clases:

Clase Georeferenciacion: crea un sistema de referenciación espacial asociado con una entidad espacial.

Clase UTM: contiene las operaciones matemáticas necesarias para realizar una proyección del sistema de coordenadas geográficas a una zona determinada del sistema UTM.

Clase OpeGrados: realiza las operaciones necesarias para la conversión de grados a grados, minutos y segundos.

Paquete JGeoST.importar

A Descripción: este paquete contiene las clases que realizan la importación de los diferentes formatos a los formatos correspondientes en ESTELLUS.

B Lista de clases:

Clase ASCII: realiza la importación de campos raster bidimensionales de formato ascii raster de ESRI a formato nativo de ESTELLUS.

Clase LeerDat: realiza la importación de entidades puntuales y campos raster desde archivos *.dat a formato nativo de ESTELLUS.

Clase ImportarASCII: es un JDialog que posibilita al usuario realizar la importación de archivos ascii raster de ESRI.

Clase ImportarDAT: es un JDialog que posibilita al usuario realizar la importación de archivos *.dat para entidades puntuales.

Clase ImportarDATRaster: es un JDialog que posibilita al usuario realizar la importación de archivos *.dat para campos raster bidimensionales y tridimensionales.

Paquete JGeoST.operaciones

A Descripción: este paquete contiene las clases que permiten realizar algunas operaciones generales ampliamente utilizadas en toda la biblioteca de clases.

B Lista de clases:

Clase Estadisticos: en esta clase se realizan las estimaciones de los estadísticos para muestras puntuales y para los campos raster.

Clase Leerxyz: lee archivos de texto en formato especial ESTELLUS para formar una lista de entidades puntuales.

Clase Guardarxyz: guarda archivos de texto en formato especial ESTELLUS que posteriormente pueden ser usados para formar una lista de entidades puntuales.

Clase ListaObservaciones: construye objetos de observaciones que pueden integrarse a los procedimientos de estimación y representan una capa vectorial con un atributo único para cada entidad contenida.

Clase Observacion: es una clase que representa a un objeto puntual en el espacio con un atributo asociado.

Clase ControlTabla: esta clase proporciona los controladores para la tabla de entrada de información puntual.

Clase DataViewEdit: es un `JDialog` que muestra la información contenida en un archivo de observaciones leído mediante la clase `Leerxyz`.

Clase EditarXYZHEntrada: es un `JFrame` que permite la entrada y edición manual de entidades puntuales del tipo `ListaObservaciones`.

Clase ModeloTabla: es una clase que implementa la interfaz `TableModel` (del paquete `swing` de Java) y permite la creación y modificación de la tabla de entrada de información.

Clase SeleccionarSecLista: es un `JDialog` que permite seleccionar un objeto secundario de las listas de observaciones de la interfaz principal de ESTELLUS.

Paquete JGeoST.raster

A Descripción: este paquete contiene las clases que manipulan los objetos raster en ESTELLUS.

B Lista de clases:

Clase Cube: construye el cubo tridimensional que se va a desplegar utilizando `VisAD`.

Clase DataRaster: manipula la información contenida en el campo raster bidimensional.

Clase DataRaster3D: manipula la información contenida en el campo raster tridimensional.

- Clase EstadisticosMapa:** lleva a cabo los cálculos de los estadísticos de los campos raster.
- Clase Histograma:** maneja el cálculo del histograma de un campo raster.
- Clase IsoCurvasRaster:** `DibujableDisplay` encargado de graficar las curvas de nivel de un campo raster bidimensional.
- Clase Matrix4D:** lleva a cabo algunas operaciones matriciales de transformación de campos.
- Clase MetaRaster:** `DibujableDisplay` que maneja los metadatos de los campos raster bidimensionales y algunas operaciones de manipulación.
- Clase MetaRaster3D:** `DibujableDisplay` que maneja los metadatos de los campos raster tridimensionales y algunas operaciones de manipulación.
- Clase TablaColores:** maneja lo referente a la paleta de colores del despliegue.
- Clase ColorManager:** `JDialog` que permite la selección de la paleta de colores del despliegue.
- Clase ColorManager3D:** `JDialog` que permite la selección de la paleta de colores del despliegue tridimensional.
- Clase DefinirIso:** `JDialog` que permite ingresar los parámetros de la extracción de las curvas de nivel del campo raster.
- Clase ImagenPaleta:** efectúa la construcción de la imagen del campo en el despliegue y la paleta de colores seleccionada.
- Clase JDialogEstadisticos:** es un diálogo que muestra los estadísticos del campo raster.
- Clase JDialogExtraerSubmapa:** es un diálogo que permite extraer un campo bidimensional subconjunto de otro existente.
- Clase JDialogExtraerSubmapa3D:** es un `JDialog` que permite extraer un campo tridimensional subconjunto de otro existente.
- Clase JDialogHistograma:** es un `JDialog` que muestra el resultado del cálculo del histograma de un campo.
- Clase JDialogHistograma:** es un `JDialog` que permite personalizar la visualización de los campos raster.
- Clase JDialogSelectorMapa:** es un `JDialog` que permite al usuario seleccionar la serie del raster a desplegar.
- Clase MapaRaster:** es un `JFrame` que despliega la visualización de los campos raster bidimensionales y contiene las utilidades que los manipulan.
- Clase Mapa3D:** es un `JFrame` que despliega la visualización de los campos raster tridimensionales y contiene las utilidades que los manipulan.

C Lista de interfaces:

- Clase Capturable:** es una clase abstracta que conceptualiza la captura de la posición del cursor del mouse sobre un despliegue de VISAD.

Clase DibujableDisplay: clase abstracta que conceptualiza el procedimiento de dibujo de una despliegue raster con VISAD.

Clase Guardable: clase abstracta que conceptualiza el procedimiento de almacenamiento en disco de un campo raster.

Paquete JGeoST.variograma

A Descripción: este paquete contiene las clases que efectúan los cálculos del variograma experimental y el ajuste de modelos.

B Lista de clases:

Clase AjustarMinimosCuadradosNelderMead: realiza el ajuste por mínimos cuadrados ponderados de un modelo de variograma a uno experimental.

Clase MapaVariograma: realiza el cálculo de un variograma superficial.

Clase VariogramaAnisotropico3D: calcula un variograma direccional de un conjunto de observaciones.

Clase VariogramaCruzadoAnisotropico3D: calcula un variograma cruzado de un conjunto de observaciones.

Clase VariogramaOmnidireccional: calcula el variograma omnidireccional de un conjunto de observaciones.

Clase VariogramaCruzadoOmnidireccional: calcula el variograma cruzado omnidireccional de un conjunto de observaciones.

Clase Vector3DOperations: realiza las operaciones tridimensionales necesarias para los cálculos del variograma experimental.

Clase Esferico: RegressionFunction de la forma de función esférica de variograma.

Clase Lineal: RegressionFunction de la forma de función lineal de variograma.

Clase Gaussiana: RegressionFunction de la forma de función gaussiana de variograma.

Clase Exponencial: RegressionFunction de la forma de función exponencial de variograma.

Clase Pepita: RegressionFunction de la forma de efecto pepita de variograma.

Clase EvaluarModeloCovarianza: efectúa la evaluación del modelo de covarianza correspondiente a una **RegressionFunction**.

Clase EvaluarModeloVariograma: efectúa la evaluación del modelo de variograma correspondiente a una **RegressionFunction**.

Clase EvaluarAnisotropia: efectúa la evaluación de un modelo anisotrópico de variograma o covarianza.

Clase SuperposicionAnisotropia: efectúa la evaluación de una superposición de modelos anisotrópicos de variograma.

Clase AjustarVariograma: JFrame que permite realizar el ajuste de variograma.

Clase Corregionalizacion: JFrame que permite realizar el diseño de la superposición de modelos anisotrópicos de variograma.

Clase VariogramaDirecto: JFrame que permite realizar el cálculo de los variogramas directos experimentales.

Clase CariogramaCruzado: JFrame que permite realizar el cálculo de los variogramas cruzados experimentales.

C Lista de Interfaces:

Interfaz RegressionFunction: es una interfaz que conceptualiza la función algebraica para ajustar a una serie de observaciones. En este caso son funciones permitidas de variograma.

Interfaz EstructuraImbricada: es una interfaz que conceptualiza una superposición lineal de `RegressionFunctions`.

Paquete JGeoST.indicador

A Descripción: este paquete contiene las clases que realizan la codificación indicadora de la información.

B Lista de clases:

Clase VariableIndicadora: realiza la codificación indicadora de una variable.

Clase VariableIndicadoraSuave: realiza la codificación indicadora de una variable secundaria.

3.5. Comentarios finales

En este capítulo se han mostrado las características generales de la librería de clases implementada. En especial, la sección 3.4 sirve para obtener una vista general de las clases desarrolladas. En los capítulos siguientes se hará una descripción completa de éstas y de los métodos más importantes incluidos, pero obviamente, debido a la extensión del código es imposible entrar a dar detalles de los métodos existente. El usuario interesado en conocer en profundidad todos los métodos, parámetros y consideraciones especiales puede remitirse al código, el cual ha sido extensamente documentado.

4

Manejo de información espacial

4.1. Introducción

La Geoestadística es una disciplina que estudia fenómenos espaciales, por lo cual la información que debe integrarse en los análisis y es usada en los procedimientos de estimación debe ser geográfica. Este tipo de información posee de forma intrínseca dos tipos de componentes (Burrough y McDonnell, 1998):

1. Una componente atributiva, la cual está formada por un conjunto de propiedades enumerables y/o medibles.
2. Una componente geográfica, la cual requiere que las entidades u objetos a representar tengan asociadas coordenadas espaciales (geográficas o planas) con referencia a un sistema que indiquen claramente su ubicación espacial.

Lo anterior implica que un análisis geoestadístico debe realizarse partiendo de información georeferenciada, con propiedades y atributos claramente establecidos, nombrables, enumerables y convenientemente cuantificados.

Bajo estas premisas, en este capítulo se describen las clases desarrolladas para el manejo de la información georeferenciada espacialmente de acuerdo con sistemas de coordenadas (bi)tridimensionales. Algunas de las clases que se describen han sido completamente desarrolladas en este trabajo, mientras que otras han surgido de adaptaciones y modificaciones de clases previamente existentes en otras librerías como HIDROSIG (Poveda et al., 2007 a). Con el objetivo de construir un entorno computacional que soporte la manipulación de la información a incluir en los procedimiento de estimación espacial y procesamiento de

los resultados, se han creado en ESTELLUS diferentes clases que manejan básicamente dos tipos de información geográfica:

- **Información puntual.** Para este tipo de información se han implementado objetos que representan entidades puntuales en formato vectorial y tienen asociadas coordenadas tridimensionales de la forma (x, y, h) o $(longitud, latitud, h)$ y un atributo que ha sido cuantificado previamente.
- **Información exhaustiva.** Para este tipo de información se han creado objetos que representan campos aleatorios en formato raster bidimensionales o tridimensionales.

Junto con las clases correspondientes al manejo de la información, se han desarrollado otras que permiten efectuar el análisis descriptivo y obtener indicadores generales del comportamiento estadístico de las muestras de la información disponible. Los resultados obtenidos después de la ejecución de los métodos contenidos en dichas clases pueden ser incorporados en análisis espaciales posteriores o pueden usarse como datos estadísticos de apoyo.

Con respecto al código de programación generado en la implementación de los objetos geográficos, en éste capítulo se describen detalladamente las características de los programas desarrollados para la manipulación de los objetos geográficos necesarios para definir un procedimiento de estimación espacial.

4.2. La referenciación espacial de las entidades geográficas

Como se dijo más arriba, Burrough y McDonnell (1998) manifiestan que una entidad o fenómeno geográfico requiere dos componentes para su correcta definición: la información como tal, es decir, sus características que pueden enumerarse, nombrarse y en algunos casos cuantificarse; y su localización espacial, que se refiere básicamente al lugar donde está ubicada dicha entidad geográfica. Un esquema conceptual del significado de la georeferenciación para un campo raster se presenta en la figura 4.1, donde se muestra un dominio de representación (Norte América, Centro América y una fracción de América del Sur) para un campo que se sitúa dentro de un contexto geográfico, en este caso bajo un sistema de coordenadas geográfico usando longitud y latitud.

Para asignar un entorno espacial a los objetos geográficos que se construyan, el cual define su donde o ubicación espacial de manera única, suficiente y clara, se dispone de la clase **Georeferenciacion**. Como paso previo a la definición del objeto computacional de campo raster, debe crearse un objeto que interprete la localización espacial del ente geográfico. Para lograr lo anterior, en este trabajo se adaptó el objeto computacional de referenciación espacial contenido en la librería HIDROSIG (Poveda et al., 2007 a,b) e implementado en su clase **Georeferenciacion**. Básicamente, el objeto de referenciación espacial contiene las características que definen el sistema de coordenadas que se usa para la representación espacial de las coordenadas asociadas con un ente geográfico. Así, éste especifica si se

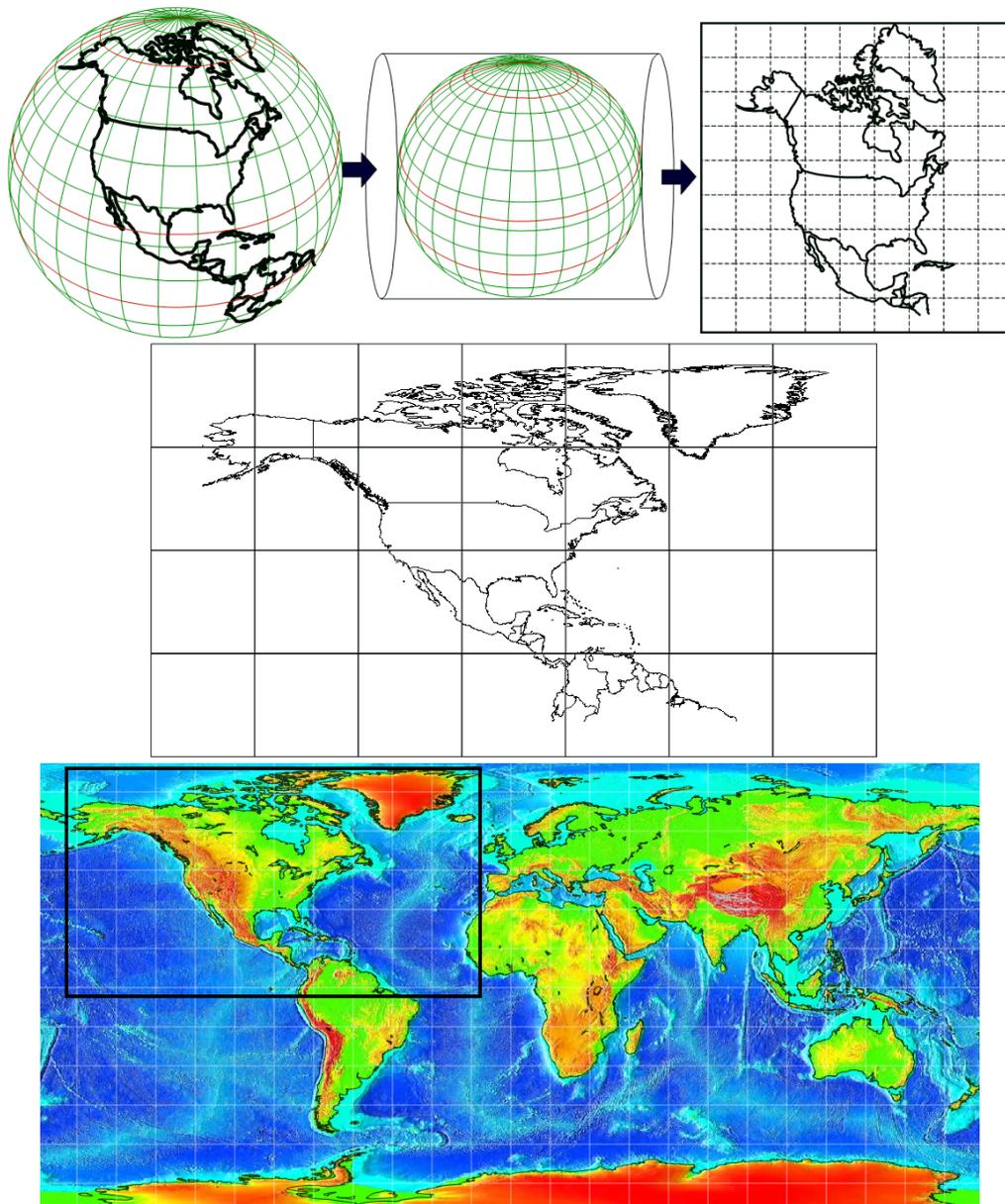


Figura 4.1. Esquema conceptual de las funciones y propiedades de los objetos pertenecientes a la clase Georeferenciación.

está trabajando con sistemas de coordenadas geográficas, en latitud y longitud o en sistemas de coordenadas planos asociados con algún tipo de proyección o si, por el contrario, el campo no tiene definido un sistema de georeferenciación estándar sino un sistema de

coordenadas locales definido a priori por el modelador. Esta última es una posibilidad inexistente en la versión original de la clase en HIDROSIG y por tanto fue implementada en el presente trabajo. Además de lo anterior, es posible realizar transformaciones de coordenadas, siempre y cuando se trabajen con los sistemas de georeferenciación soportados por la clase `Georeferenciacion`, los cuales son:

- Coordenadas geográficas para los elipsoides Internacional de Hayford y el elipsoide del datum *WGS-84*. El primero es utilizado como datum en algunas regiones del mundo, por ejemplo en Suramérica (Fernández, 2002; IGAC, 2004), mientras que el segundo es ampliamente utilizado puesto que soporta el sistema de coordenadas UTM y el sistema de posicionamiento global, *GPS* (Kennedy, 2000).
- Coordenadas planas obtenidas mediante el uso de la proyección Universal Transversa de Mercator, UTM. Los detalles matemáticos y las características generales de este sistema de proyección pueden consultarse en los textos de Fernández-Coppel (sin fecha), Martín-Asín (1990), Fernández (2002) y Snyder (1987). Detalles operacionales para *SIG* comerciales son presentados por Kennedy (2000).
- Algunos sistemas de coordenadas utilizados para la representación cartográfica en el territorio colombiano. Los lectores interesados en los detalles de estos sistemas pueden remitirse a los textos de Fernández (2002) e IGAC (2004).

El objeto de la clase `Georeferenciacion` se construye definiendo los valores de los parámetros del elipsoide asociados con la referenciación espacial geográfica y las posibles proyecciones que puedan llevarse a cabo. Dichos parámetros son: los tamaños de los tres semiejes del elipsoide, el nombre del elipsoide, el factor de escala, las coordenadas geográficas de origen y el falso origen. Estos parámetros se usan para, de ser necesario, aplicar transformaciones sobre el dominio espacial representado en el campo raster o las coordenadas asociadas con las entidades puntuales. Cuando se trata de campos tridimensionales, las transformaciones de sistemas de coordenadas previamente comentados se realizan sólo en las correspondientes latitud y longitud o norte (*Y*) y este (*X*), mientras que las coordenadas que representan la profundidad no se modifican. Todos los objetos computacionales geográficos (los cuales se presentan en las siguientes secciones de este capítulo) tienen un objeto de la clase `Georeferenciacion` como parámetro, puesto que éste representa su componente espacial, es decir, contiene los detalles asociados con su localización y el tipo de coordenadas que se va a utilizar. La clase `Georeferenciacion` se localiza en el paquete `JGeoST.georef`. objetos de la librería de clases de ESTELLUS.

4.3. La manipulación de información puntual

Usualmente, la información disponible en un análisis geoestadístico está localizada de forma discreta en el espacio. En este caso cada una de las localizaciones de los datos disponibles puede asociarse con entidades puntuales espaciales. Por ejemplo, fuentes de

información puntual frecuentemente empleadas son registros de estaciones climatológicas que tienen asociadas mediciones de variables atmosféricas como temperatura, presión atmosférica, brillo solar, precipitación, etc.; perforaciones en el terreno en las cuales se registra la altura piezométrica y/o la concentración de un contaminante en el subsuelo. En estos casos se puede considerar a la localización del muestreo como una entidad geográfica puntual, a la cual se asocian sus coordenadas espaciales (x_i, y_i, h_i) y una cantidad de atributos Z_i , generalmente relacionados con las mediciones realizadas localmente. En ESTELLUS se han implementado dos clases Java que hacen posible el manejo de los atributos y las coordenadas asociadas con entidades geográficas puntuales, que se han denominado `Observacion` y `ListaObservaciones`.

4.3.1. La clase `Observacion`

Una observación se refiere simplemente a una de las múltiples realizaciones disponibles de la variable que se desea analizar en una zona de estudio, a la cual se le asocia el valor de un atributo que ha sido cuantificado o medido. La implementación computacional de este tipo de objetos se ha realizado en la clase `Observacion`, localizada en el paquete `JGeosST.operaciones.objetos`. Su constructor recibe como parámetros las coordenadas (x, y, h) (este, norte, profundidad o longitud, latitud, profundidad) que definen la ubicación de la localización y el valor de la observación de la variable disponible en ese punto, Z . Los métodos implementados en esta clase hacen posible la designación de atributos y coordenadas asociadas con la localización y la variable representada y la extracción automática de éstos.

4.3.2. La clase `ListaObservaciones`

Cuando se desea hacer un análisis del comportamiento y/o estimación espacial de una variable debe disponerse de un conjunto de observaciones en el dominio de trabajo. Para ello se ha implementado un objeto computacional que hace posible el manejo de un conjunto de observaciones en la clase `ListaObservaciones`, contenida en el paquete `JGeoST.operaciones.objetos`. Cuando se construye un objeto del tipo `ListaObservaciones`, se crea un arreglo de objetos de la clase `Observacion` a partir de una matriz que contiene la localización de todos los puntos geográficos en donde se dispone de información y el valor de las observaciones de un atributo asociado. La información asociada a cada una de las observaciones puede extraerse usando métodos de clase, de forma que está disponible para ser incluida en los análisis implementados. Los siguientes son los parámetros necesarios para construir un objeto de esta clase:

- **Matriz de localizaciones:** es una matriz en la cual se almacenan las coordenadas de las localizaciones para las cuales se define información del atributo analizado. Cada columna de la matriz está formada por un vector que cumple estrictamente con la forma $[este, norte, prof.]$ o $[longitud, latitud, prof.]$.

- **Vector de observaciones:** en este vector se almacenan las realizaciones de la variable o atributo de análisis en las localizaciones disponibles. Cada componente se asocia con la misma columna componente en la matriz de localizaciones.
- **Nombre de la variable:** es una cadena de caracteres que representa el nombre de la variable analizada.
- **Georeferenciación:** es un objeto de la clase `Georeferenciacion` que define la ubicación de los puntos de acuerdo con un sistema de coordenadas estándar definido a priori o un sistema plano local construido por el analista.

Adicionalmente, la clase `ListaObservaciones` dispone de métodos para el manejo de la información contenida en el arreglo de objetos de la clase `Observacion`, los cuales pueden ser utilizados para redefinir y extraer coordenadas y valores de las localizaciones de las realizaciones. Más aún, al recibir como parámetro un objeto de la clase `Georeferenciacion`, es posible realizar transformaciones de coordenadas y proyecciones geográficas sobre las coordenadas de los puntos.

4.4. La manipulación de información exhaustivamente conocida

En ESTELLUS la información de salida de los algoritmos de estimación y la información exhaustivamente muestreada se representa mediante campos en formato raster. Las características generales de este tipo de representación de información espacial pueden encontrarse en cualquier texto básico de introducción a los sistemas de información geográfica (Meijernik et al., 1994; Bosque-Sendra, 1997; Burrough y McDonnell, 1998). Una descripción de campo raster ampliamente aceptada es la propuesta por Burrough y McDonnell (1998), quienes consideran que el formato raster es el más apropiado para la manipulación de información que varía continuamente en el espacio, mediante la división de un dominio espacial de análisis en pequeños subdominios, denominados *pixels* bidimensionales o *voxels* tridimensionales, para cada uno de los cuales las propiedades de la variable representada son constantes en su interior.

Así, los objetos implementados para la representación de campos raster incluyen tanto parámetros para la definición de su dominio espacial, como parámetros para la manipulación de la información de la variable contenida dentro de éste. Las clases de Java implementadas para estos propósitos son modificaciones de clases originalmente incluidas en HIDROSIG (Poveda et al., 2007 a,b), que actualmente sólo soportan objetos raster bidimensionales. En este trabajo se realizó una extensión de dichos objetos de forma que sea posible la manipulación computacional de campos raster tridimensionales. A continuación se presenta una descripción de dichos objetos y las clases Java implementadas para su manipulación.

4.4.1. Las clases `DataRaster` y `DataRaster3D`

De acuerdo con lo sugerido por Burrough y McDonnell (1998) y comentado más arriba, además de su localización, una entidad geográfica también está caracterizada por su información y propiedades asociadas. En el caso de entidades geográficas representadas mediante campos raster sólo puede asociarse una característica o atributo, el cual es variable en el espacio. La manera usual de representar la información raster es mediante el modelo matricial rectangular (Bosque-Sendra, 1997), en el cual cada celda del campo tiene asociado un valor del atributo representado. Esto, inicialmente, genera problemas de eficiencia computacional puesto que se requiere de un gran recurso de memoria para almacenar la información contenida. A pesar de ello, cuando se trabaja con propiedades como la permeabilidad, la porosidad, la temperatura, la precipitación, etc., se hace necesario su implementación puesto que ellas son altamente fluctuantes sobre regiones muy pequeñas. En general, existen modelos de almacenamiento raster más elaborados y eficientes, pero tienen la gran desventaja de estar diseñados para efectuar el manejo de información que es constante por sub-regiones consideradas homogéneas, contenidas dentro del dominio completo del campo (Burrough y McDonnell, 1998). Considerando lo anterior, la representación de campos raster en ESTELLUS se lleva a cabo utilizando el modelo matricial rectangular. Para lograrlo se adaptaron las clases que crean y realizan la manipulación computacional de los campos raster (bidimensionales) existentes en la librería de HIDROSIG, y se extendió su alcance a dominios espaciales tridimensionales. La eficiencia de estas clases para efectos de implementación de funciones de manipulación raster ha sido evidenciada en diferentes aplicaciones como las presentadas por Barco y Cuartas (1998), UPME (2000), Vélez et al. (2000), Ramírez (2002), Álvarez y Alarcón (2004), Álvarez (2007) y Poveda et al. (2007 b).

En HIDROSIG, la información atribuida al dominio del raster está compuesta por un conjunto de campos que pueden representar tanto mapas resultantes del procedimiento de corrección de un modelo digital de elevaciones o campos con información correspondiente a series mensuales de una variable hidrológica, por ejemplo precipitación, temperatura, brillo solar, etc. (ver figura 4.2). Así, el objeto de manipulación de información raster efectúa las operaciones sobre series de campos y no sobre un campo único (Poveda et al., 2007 a,b). Lo anterior resulta bastante útil y se incorpora en ESTELLUS con el fin de representar campos resultantes de un procedimiento de estimación mediante krigeado que incluyen los valores del atributo estimados para cada *pixel*, junto con su varianza y desviación estándar asociadas. Esto también puede ser de ayuda en el futuro cuando se implemente la modelación de funciones locales acumuladas de probabilidad, para representar los campos de probabilidades de una forma muy compacta y hacer posible la creación automática de algoritmos de simulación secuencial.

Las clases `DataRaster` y `DataRaster3D`, dentro del paquete `JGeoST.raster.objetos`, requieren la definición de tres parámetros:

- **Vector de información:** está compuesto por el conjunto de datos que se desea representar en el campo raster, su tamaño debe ser el producto del número de filas,

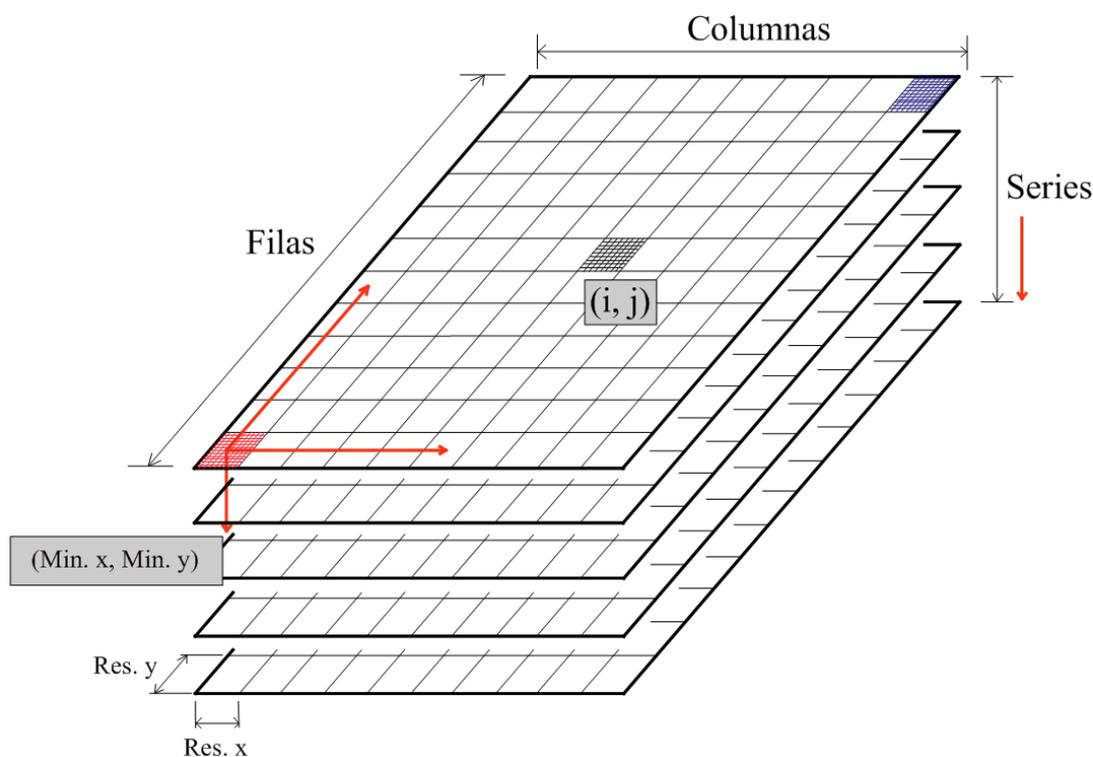


Figura 4.2. Esquema conceptual empleado para la representación del dominio espacial de un campo raster bidimensional, implementado en la clase `MetaRaster`.

columnas y niveles del campo raster a representar. La manipulación de la información y el tamaño computacional en memoria RAM del objeto está determinado por el formato del número: entero, flotante, doble, byte, etc. En los *pixels* donde no se dispone de información se debe especificar un valor denominado *faltante*, el cual puede representarse mediante una sentencia *Not a Number* (NaN) o un valor que no es posible obtenerse físicamente, por ejemplo -9999.99 .

- **Serie:** es un entero corto codificado en formato byte que indica la posición del campo actual dentro de la serie de campos que conforman la información completa del raster. En el caso particular de ESTELLUS, se define como 0 para el campo de valores estimados, 1 para el campo de varianzas asociadas y 2 para el campo de desviaciones estándar. Las tres series de información tienen asociado el mismo dominio espacial y metadatos del campo raster como se describe en la próxima sección.

- **Nombre:** es una cadena de caracteres que representa el nombre asociado con el vector de datos que se ingresa a la serie de vectores, es decir, el nombre correspondiente al vector definido en la posición serie.

En ambas clases existen métodos de cálculo denominados `getValor`, los cuales permiten la extracción de los valores del atributo representado en el campo raster para una posición interna de la matriz de datos, y son muy útiles en el momento de extraer la información secundaria *pixel* por *pixel* para la ejecución de la estimación mediante cokrigado colocalizado o krigado con deriva externa.

4.4.2. Las clases `MetaRaster` y `MetaRaster3D`

En ESTELLUS la representación de campos raster, junto con la manipulación de su información y la definición de su dominio espacial se realiza teniendo en cuenta:

1. Los objetos computacionales definidos en las clases `DataRaster(3D)`, que permiten la extracción y modificación de la información de la variable representada para una determinada posición en la serie de campos raster considerados.
2. Un objeto de la clase `Georeferenciacion`, que contiene la información acerca del sistema de coordenadas usado para representar el dominio espacial del campo.
3. Algunas características acerca de la localización espacial del campo raster, como pueden ser las coordenadas de uno o varios puntos de la malla; para lo cual, en general, se usan las coordenadas de una de sus esquinas.
4. Las características de la malla que definen la geometría del dominio espacial representado por el campo raster, como lo son la resolución espacial o tamaño espacial de celdas (*pixels* o *voxels*) y el número total de celdas, para cada dirección de ejes coordenados.

Los constructores y métodos que crean las propiedades y manipulan la información de los objetos computacionales para los campos raster se han programado en las clases `MetaRaster` y `MetaRaster3D`. Éstas han sido el resultado de efectuar un procedimiento de adaptación y revisión de una clase originalmente denominada `MetaRaster`, que está contenida dentro de la biblioteca de clases de HIDROSIG. Su alcance se ha ampliado para lograr que los objetos raster en ESTELLUS puedan afrontar problemas de análisis y estimación tridimensional. Las clases implementadas para la manipulación raster, `MetaRaster` y `MetaRaster3D`, pertenecen al paquete `JGeoST.raster.objetos` y los parámetros que deben definirse para su creación son:

- **Coordenadas inferiores del campo:** se trata de las coordenadas de la celda inferior izquierda del campo raster. Representan un origen de coordenadas internas para el dominio espacial representado (ver figuras 4.2 y 4.3), las cuales deben ser consecuentes con el sistema de georeferenciación adoptado. Se definen entonces tres

coordenadas, dos en los ejes latitud-longitud o este-norte y una en la dirección profundidad o elevación, de ser requerido.

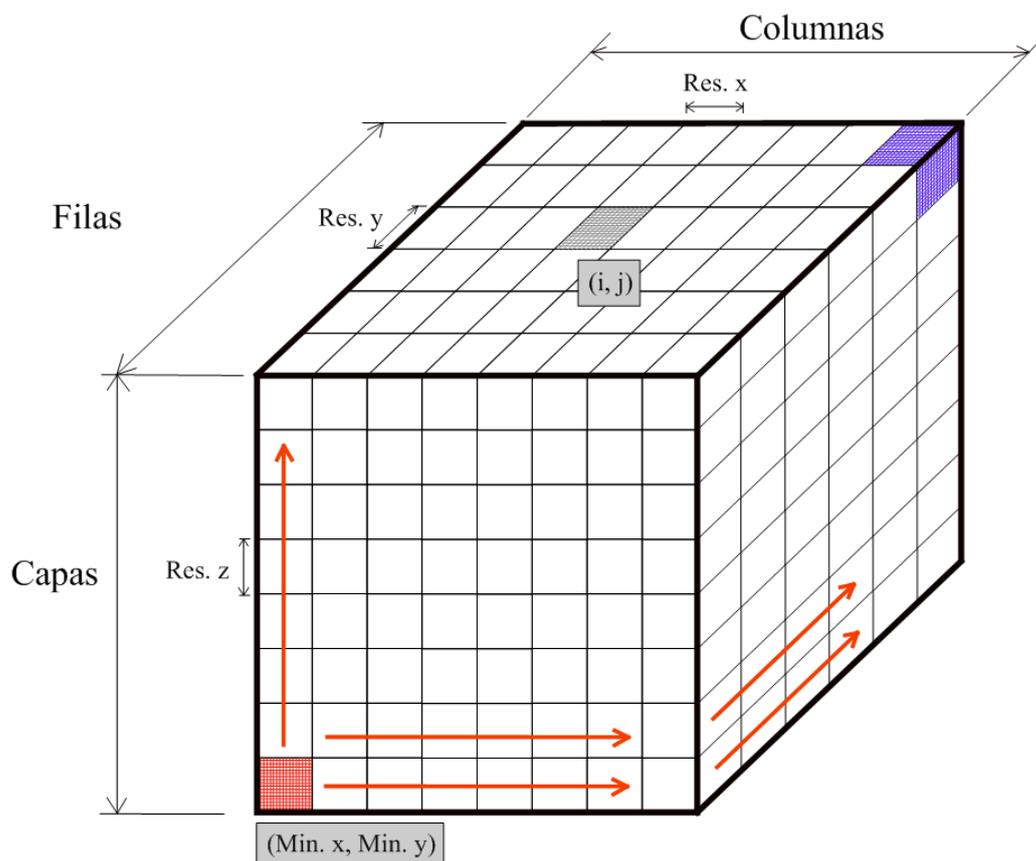


Figura 4.3. Esquema conceptual empleado para la representación del dominios espacial de un campo raster tridimensional, implementado en la clase `MetaRaster3D`.

- **Tamaño de celda:** se refiere a los diferentes tamaños espaciales adoptados en las direcciones de los ejes de georeferenciación, es decir, el tamaño de cada lado de la celda en dirección de los ejes longitud-latitud-elevación o este-norte-elevación (ver figuras 4.2 y 4.3). El tamaño de la celda está relacionado con la resolución del campo, es decir, con el tamaño de la mínima entidad espacial que puede representarse en el campo, lo cual, a su vez, representa la escala del raster.
- **Número de celdas:** son la cantidad de celdas que forman la matriz de datos del campo raster, los cuales se definen por direcciones: el número de filas se define en dirección latitudinal o norte, el número de columnas se define en dirección longitu-

dinal o este y el número de niveles se define en dirección de la elevación (ver figuras 4.2 y 4.3).

- **Formato del número:** el formato define las características de los números empleados en los cálculos y representados en el campo raster. Tiene que ver con la forma virtual como el ordenador entiende y almacena la información contenida en el número que representa el valor de la variable representada, celda a celda. Su correcta elección es importante puesto que, en algunos casos, cierto tipo de formato de número puede no llegar a ser conveniente ya sea por la variable representada o por el tamaño de memoria de ordenador requerida para el manejo del campo raster. Por ejemplo, un modelo digital de elevaciones a escala regional puede ser bien representado mediante números enteros, puesto que representa una solución de compromiso entre el detalle al cual se quiere representar la variable y los requerimientos computacionales. En contraposición, la misma elección puede no ser conveniente para el campo de una variable cuyo rango de variación es pequeño, como la temperatura en una región. Además, las unidades de representación de la variable pueden llegar a modificar la validez de la adopción de cierto tipo de formatos de número. En el caso de este trabajo, la adopción de un cierto formato de número afecta los métodos que actúan sobre un objeto de las clases `DataRaster` y `DataRaster3D`, mejorando o empeorando la velocidad y requerimientos de memoria en la ejecución de los algoritmos.
- **Faltante:** es el valor que debe ubicarse en las componentes de los vectores de datos que correspondan a localizaciones en el campo raster donde no se dispone de información.
- **Georeferenciación:** es un objeto de la clase `Georeferenciacion` que representa la georeferenciación espacial del campo raster y hace posible que sobre éste puedan actuar los métodos de proyección geodésica incluidos en dicha clase.
- **DataRaster asociados:** es un arreglo de objetos de las clases `DataRaster` o `DataRaster3D`, que asocia la información disponible al campo raster y activa su manipulación mediante su asociación con las características espaciales del raster. Es decir, asocia las coordenadas locales de matriz del vector de información con las coordenadas espaciales del sistema de georeferenciación del campo.
- **Otros parámetros:** aparte de los anteriores, deben designarse otros parámetros en forma de cadenas de caracteres: su nombre, la unidad en la cual se está representando la variable asociada, el nombre de la variable representada, comentarios, nombre de la región representada, etc. Estos parámetros no intervienen en las rutinas requeridas por los algoritmos de manipulación raster, pero se usan para generar encabezados de texto al momento de utilizar el raster en la interfaz principal de la aplicación.

En ambas clases, `MetaRaster` y `MetaRaster3D`, se han adaptado los métodos existentes en la librería de `HIDROSIG` para la manipulación y análisis de los campos raster bidimensionales y su post-procesamiento, de ser necesario, como lo son el cálculo del histograma

del mapa, la extracción de un campo contenido dentro del original y el almacenamiento del raster en un disco.

Cálculo del histograma de un campo raster

Dentro del catálogo de algoritmos de manipulación de campos raster de ESTELLUS se ha implementado un método que permite extraer de forma automática los histogramas de los campos raster (bi)tridimensionales. El método recorre cada una de las celdas que componen el campo y, de acuerdo con el número de clases pre-establecido, clasifica los valores del atributo representado. Este cálculo permite, como se mostrará en el próximo capítulo, la reclasificación por atributo interactiva del campo analizado por parte del usuario.

Extracción de campos

En algunas ocasiones se hace necesario extraer un campo raster que es un subconjunto del campo original. Para la librería ESTELLUS se ha escrito un método de extracción de sub-campo tridimensional que es una extensión del algoritmo existente en la librería de HIDROSIG, el cual está operativo para dominios espaciales (bi)tridimensionales. El algoritmo recibe como entrada las coordenadas de las esquinas inferior izquierda y superior derecha que definen un cubo en dominio tridimensionales. A partir de ellas recorre el dominio del campo original y detecta que celdas se encuentran dentro del dominio deseado para asignar los valores correspondientes al nuevo vector de datos destinado para el manejo de la información del nuevo campo raster. Esto se realiza para el campo activo de la serie de rasters en el objeto de `DataRaster(3D)`. El método arroja como salida un nuevo objeto de la clase `MetaRaster` o `MetaRaster3D` que corresponde al conjunto espacial que se desea extraer.

Almacenamiento del campo en el disco

Mediante la adaptación de los métodos de almacenamiento de campos raster existentes en HIDROSIG se han implementado nuevos métodos que permiten el almacenamiento de objetos computacionales pertenecientes a las clases `MetaRaster` y `MetaRaster3D`. El almacenamiento consiste en la creación de cuatro archivos. El primero es de texto y almacena las características o parámetros de los objetos de las clases `MetaRaster` y `MetaRaster3D`. Los archivos generados presentan la extensión `*.MetaGEO` y `*.Meta3GEO`, y pueden abrirse directamente con cualquier editor convencional de texto. Los tres restantes, con extensión `*.GEO`, almacenan los vectores de datos para las series de valores, es decir, existe un archivo que almacena el vector de datos del campo de valores estimados, otro que almacena el vector de varianzas y un último archivo que almacena el vector de desviaciones estándar. La información contenida en estos tres archivos está codificada en el formato binario nativo de Java, por lo cual sólo es posible leerse creando un programa en dicho lenguaje que lea el formato establecido. El formato usado es bastante simple: en cada archivo la información

se almacena desde la celda inferior izquierda del cubo de domino espacial del campo hasta la celda superior derecha, siguiendo una ruta que recorre jerárquicamente, primero las columnas de una fila, posteriormente las filas de cada nivel y finalmente repite el recorrido nivel por nivel hasta llegar a la celda destino. En las figuras 4.2 y 4.3 se muestran flechas de color rojo que indican la trayectoria seguida por el algoritmo para el almacenamiento de la información.

4.5. Algoritmos para el análisis descriptivo de los datos

Como señalan diferentes investigadores (Isaaks y Srivastava, 1989; Pebesma y Wesseling, 1998; Carr y Mela, 1998; Wingle et al., 1999; Getis, 1999), el análisis descriptivo de la información permite organizar los datos y proporciona indicios acerca de las características estadísticas de los mismos. Los algoritmos implementados y/o adaptados en este trabajo se dividen en univariados y bivariados según se analice una o dos variables.

4.5.1. Análisis univariado

Los algoritmos de análisis univariado implementados en la clase `DibujarHistograma`, pertenecen al paquete `JGeoST.descriptiva.widgets`, y son los siguientes:

- **Cálculo de histogramas:** para la estimación de los histogramas se ha adaptado el método existente en la librería `JFreeChart`. La versión original sólo estima los histogramas, mientras que la versión adaptada estima tanto los histogramas acumulados como los histogramas acumulados relativos. Los parámetros requeridos por el algoritmo son: un vector con la serie de datos, los límites superior e inferior de la variable y el tamaño de la clase, y a partir de éstos estima los tres tipos de histograma previamente comentados. El lector interesado en conocer en mayor profundidad el API de la librería `JFreeChart` puede remitirse a la página web oficial del proyecto: www.jfree.org/jfreechart/. En especial, para los cálculos del histograma debe remitirse a la clase `HistogramDataSet`.
- **Cálculo de estadísticos:** el cálculo de los estadísticos se realiza mediante la utilización de la clase `Stat` perteneciente a la librería científica y numérica creada por el Dr. Flanagan (2008), como se comentó previamente.

4.5.2. Análisis bivariado

Los algoritmos de análisis bivariado implementados en la clase `DibujarBivariados`, pertenecen al paquete `JGeoST.descriptiva.widgets`, y son los siguientes:

- **Cálculo de estadísticos:** Para realizar el análisis bivariado se han utilizado algoritmos incluidos en la clase `Stat` de la librería científica y numérica del profesor Flanagan (2008). Se calculan el coeficiente de correlación de Pearson, la covarianza y el momento de inercia.

- **Diagramas de dispersión:** Para la estimación de los diagramas de dispersión se ha utilizado la clase correspondiente de la librería `JFreeChart`, denominada `ScatterPlot`.

4.6. Codificación indicadora de la información

Para aquellas técnicas que trabajan con el concepto de variable indicadora (Journel, 1983), se han implementado una serie de algoritmos que procesan la información original y la preparan para ser tratada en el ámbito indicador.

4.6.1. La clase `VariableIndicadora`

La clase `VariableIndicadora` ha sido diseñada para crear los objetos computacionales que estiman y operan sobre variables indicadoras. Los parámetros que recibe el constructor de la clase para cada valor de la variable para el cual se quiera calcular la transformada indicadora son:

- **Vector de observaciones:** es un arreglo con los valores de las realizaciones de la variable principal de análisis, las cuales pueden extraerse de un objeto de la clase `ListaObservaciones`. Éste tiene como propiedades las coordenadas de las entidades geográficas puntuales, el valor de una variable en dichos puntos y el sistema de georeferenciación asociado con las coordenadas. Además, sobre él pueden aplicarse los métodos de extracción de información de la lista, que en este caso particular corresponden a un arreglo con los valores del atributo de la variable original.
- **El valor del umbral:** es un valor de la variable principal que se utiliza para la transformación indicadora de la variable.

A partir del conocimiento de estos parámetros, el objeto puede someterse al método `getIndicadores`, el cual aplica la expresión de la variable indicadora para estimar la variable transformada.

4.6.2. La clase `VariableIndicadoraSuave`

La clase `VariableIndicadoraSuave` ha sido creada para incorporar en el proceso de estimación información blanda, esto es, para la implementación del cokrigado indicador. Los parámetros necesarios para construir un objeto computacional de la clase `VariableIndicadoraSuave` son:

- **Observaciones de la variable principal:** es un vector que contiene los valores observados de la variable principal.
- **Observaciones de la variable secundaria:** es un vector que contiene los valores observados de la variable secundaria en las localizaciones donde se conoce la variable principal.

- **El valor del umbral para la variable principal:** es un valor de la variable principal que se determina a priori para examinar la no excedencia de los valores en la muestra de observaciones de la variable principal.
- **Vector con los umbrales de clase para la variable secundaria:** es un vector que contiene los límites de clase de la variable secundaria de acuerdo con un número de clases definido a priori para calcular la distribución de probabilidades previa de la variable secundaria.

Para obtener los valores de la transformación para la variable secundaria se usa el método `getIndicadoresSec`, que devuelve un vector con el valor de la transformación correspondiente.

4.7. Comentarios finales

Las clases que se han discutido y descrito a lo largo de este capítulo contienen los procedimientos para representar los objetos geográficos en los cuales se sustentan los procesos de modelación del variograma y estimación espacial y pueden considerarse como el corazón de ESTELLUS. Aparte de esto se han presentado diferentes operaciones de manipulación y análisis de la información geográfica contenida en los objetos computacionales creados y/o adaptados a tal fin.

5

Análisis estructural

5.1. Introducción

El análisis estructural se lleva a cabo para caracterizar la continuidad espacial del fenómeno estudiado. Su consecución permite obtener un modelo matemático del comportamiento espacial del o de los parámetros que se registran para conocer como se comporta un sistema bajo estudio. En otras palabras, realizar un análisis estructural consiste en estimar una función matemática que refleje la correlación espacial de las variables que se observan. Las funciones matemáticas utilizadas más frecuentemente, pero no las únicas, son la función de covarianza, el correlograma y el variograma (Isaaks y Srivastava, 1989; Olea, 1991; Goovaerts, 1997; Burrough y McDonnell, 1998). A lo largo de este capítulo se presentan las bases conceptuales de los algoritmos de análisis estructural implementados en ESTELLUS y se hace una discusión extensa acerca de las clases `Java` que componen la librería con énfasis en la modelación de variogramas.

5.2. Estimación del variograma experimental directo

Dada una función aleatoria $Z(\mathbf{u})$, el variograma $\gamma(\mathbf{h})$ se define como la mitad del valor esperado de las diferencias al cuadrado entre variables aleatorias separadas por una distancia \mathbf{h} , esto es:

$$\gamma(\mathbf{h}) = \frac{1}{2}E\{[Z(\mathbf{u}) - Z(\mathbf{u} + \mathbf{h})]^2\} \quad (5.1)$$

donde $Z(\mathbf{u})$ y $Z(\mathbf{u} + \mathbf{h})$ son dos variables aleatorias separadas \mathbf{h} .

En la práctica, a partir de un conjunto de datos disponibles (realizaciones de la función aleatoria $Z(\mathbf{u})$), puede utilizarse el siguiente estimador medio para calcular el valor del variograma:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2n(\mathbf{h})} \sum_{\alpha=1}^{n(\mathbf{h})} [z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h})]^2 \quad (5.2)$$

donde $[z(\mathbf{u}_\alpha) - z(\mathbf{u}_\alpha + \mathbf{h})]$ es el incremento del atributo Z para un vector separación \mathbf{h} y $n(\mathbf{h})$ es el número de observaciones encontradas para esa distancia.

5.3. Estimación del variograma experimental cruzado

Dadas dos funciones aleatorias $Z_i(\mathbf{u})$ y $Z_j(\mathbf{u})$, el variograma cruzado se define como:

$$\gamma_{ij}(\mathbf{h}) = \frac{1}{2} E\{[Z_i(\mathbf{u}) - Z_i(\mathbf{u} + \mathbf{h})][Z_j(\mathbf{u}) - Z_j(\mathbf{u} + \mathbf{h})]\} \quad (5.3)$$

donde por un lado $Z_i(\mathbf{u})$ y $Z_i(\mathbf{u} + \mathbf{h})$ y por otro $Z_j(\mathbf{u})$ y $Z_j(\mathbf{u} + \mathbf{h})$ son pares de dos variables aleatorias separadas \mathbf{h} .

Como para el caso del variograma directo, en la práctica se utiliza un estimador medio como el siguiente:

$$\hat{\gamma}_{ij}(\mathbf{h}) = \frac{1}{2n(\mathbf{h})} \sum_{\alpha=1}^{n(\mathbf{h})} [z_i(\mathbf{u}_\alpha) - z_i(\mathbf{u}_\alpha + \mathbf{h})][z_j(\mathbf{u}_\alpha) - z_j(\mathbf{u}_\alpha + \mathbf{h})] \quad (5.4)$$

donde $[z_i(\mathbf{u}_\alpha) - z_i(\mathbf{u}_\alpha + \mathbf{h})]$ y $[z_j(\mathbf{u}_\alpha) - z_j(\mathbf{u}_\alpha + \mathbf{h})]$ son incrementos de los atributos Z_i y Z_j para un vector separación $\|\mathbf{h}\|$ y $n(\mathbf{h})$ es el número de observaciones encontradas para esa distancia.

5.4. Parámetros para el cálculo de variogramas experimentales

Cuando los datos disponibles se encuentran irregularmente distribuidos en el espacio, la aplicación de las expresiones 5.2 o 5.4 requiere de implementar una serie de tolerancias para que los estimadores obtenidos sean representativos. En la figura 5.1 puede verse un esquema de la situación mencionada. La asociación de parejas de valores de las variables se realiza teniendo en cuenta una serie de tolerancias en el vector separación, tanto desde el punto de vista de su dirección como de su módulo. En primer lugar se establece una tolerancia en el módulo que en este caso no se fija por el usuario y que se considera igual a la mitad del incremento de \mathbf{h} o *lag*. En segundo lugar se establece una tolerancia en la dirección θ considerada que se denota como $\Delta\theta$ y por último una tolerancia perpendicular a la dirección θ llamada ancho de banda b .

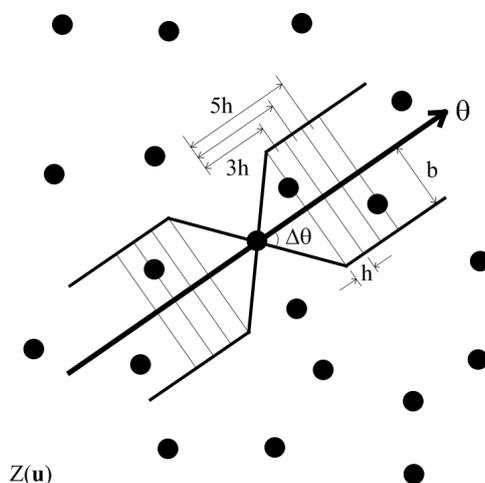


Figura 5.1. Esquema de algunos de los parámetros relativos al cálculo de variogramas experimentales para datos irregularmente distribuidos en el espacio.

5.5. Algoritmos para el cálculo de variogramas experimentales

Para el cálculo de variogramas experimentales de datos irregularmente distribuidos en el espacio se han creado cuatro objetos computacionales implementados en las cuatro clases de Java siguientes:

- La clase `VariogramaOmnidireccional` para el cálculo de variogramas omnidireccionales directos.
- La clase `VariogramaCruzadoOmnidireccional` para el cálculo de variogramas omnidireccionales cruzados.
- La clase `VariogramaAnisotropico3D` para el cálculo de variogramas direccionales directos.
- La clase `VariogramaCruzadoAnisotropico3D` para el cálculo de variogramas direccionales cruzados.

Todas ellas se hallan incluidas en el paquete `JGeoST.variograma.objetos.experimental` y se describen a continuación.

5.5.1. Variogramas omnidireccionales

Un objeto para variograma omnidireccional (directo o cruzado), creado en las clases `VariogramaOmnidireccional` y `VariogramaCruzadoOmnidireccional`, tienen la propiedad

de que el variograma omnidireccional puede calcularse empleando los respectivos métodos de clase a partir de las coordenadas de las localizaciones de las realizaciones y los valores correspondientes para la variable de análisis, junto con un tamaño del *lag* definido a priori. Los parámetros comunes de los constructores en ambas clases son:

- **Coordenadas de las localizaciones:** son las coordenadas de la ubicación de las realizaciones disponibles. Para los constructores correspondientes a los objetos diseñados para estimar variogramas cruzados también se definen las coordenadas para las realizaciones de la variable secundaria. Las coordenadas están sujetas a un sistema de georeferenciación estándar definido a priori, o simplemente, pueden ser coordenadas referidas a un sistema ficticio local. Las coordenadas de las localizaciones pueden extraerse de un objeto de la clase `ListaObservaciones` utilizando los métodos de clase.
- **Valores de las realizaciones:** son los valores de las observaciones disponibles para la variable principal y/o la variable secundaria si se trata de un análisis de variograma cruzado. Los valores de las realizaciones pueden extraerse de un objeto previamente construido perteneciente a la clase `ListaObservaciones` empleando los métodos asociados. Igualmente, este vector puede ser ingresado directamente como parámetro de clase y debe estar ordenado de forma que cada valor corresponda a las coordenadas de cada columna en la matriz de localizaciones.
- **Tamaño de lag:** es la distancia que tiene asociada cada una de las clases del variograma referida a un sistema de coordenadas planas.
- **Sistema de georeferenciación de las localizaciones:** es un objeto de la clase `Georeferenciacion` que define el sistema de referenciación espacial para las coordenadas. En el caso de realizarse un análisis de variograma cruzado debe definirse la georeferenciación de las localizaciones de la variable secundaria. Para ambos casos, si se trata de un sistema de coordenadas geográfico, las coordenadas se proyectan a un sistema de coordenadas plano soportado por la aplicación. Esto se lleva a cabo para disminuir los errores asociados con el cálculo de la distancia euclídea que se presentan al realizar la estimación de magnitudes vectoriales sobre superficies curvas, como lo son los elipsoides de referencia para los datums geográficos, en regiones de extensión considerable.
- **Sistema de georeferenciación destino:** es un objeto de clase `Georeferenciacion` que define el sistema de coordenadas al cual se va a proyectar las coordenadas de las localizaciones para propósitos de cálculo de distancias euclídeas.

Los constructores de ambas clases realizan la labor de obtener un conjunto de coordenadas de localizaciones en un sistema de georeferenciación plano. El cálculo del variograma es llevado a cabo por los métodos `calcularVariogramaOmnidireccional` y, alternativamente en el caso cruzado, `calcularVariogramaCruzadoOmnidireccional`, en los cuales

se utilizan las expresiones clásicas de estimación de variogramas directos y cruzados presentados en las ecuaciones 5.2 y 5.4. En ambas clases se han creado métodos para guardar y leer variogramas omnidireccionales, nombrados como `leerVariogramaOmnidireccional` y `guardarVariogramaOmnidireccional`. La entrada y la salida de estos métodos es un archivo de texto en el cual se escribe la distancia y su correspondiente valor del variograma. Finalmente, si se desean estimar los variogramas omnidireccionales directos de un conjunto de variables indicadoras asociados a una lista de localizaciones, basta con usar los constructores y métodos correspondientes cambiando el parámetro de vector de observaciones por el vector de transformadas indicadoras.

5.5.2. Variogramas direccionales

Los objetos de variograma direccional directo y cruzado, creados computacionalmente en las clases `VariogramaAnisotropico3D` y `VariogramaCruzadoAnisotropico3D`, ubicadas dentro del paquete `JGeoST.variograma.objetos.experimental`, se construyen a partir de los mismos parámetros que para los omnidireccionales más los dos siguientes:

- **Tolerancia angular:** es el ángulo máximo que define la zona de búsqueda para el variograma direccional, $\Delta\theta$ en la figura 5.1.
- **Ancho de banda:** es el ancho, en términos de distancia euclídea, que tiene la zona de búsqueda, b en la figura 5.1.

El constructor del objeto de variograma direccional, cruzado o directo, implementado en las clases `VariogramaAnisotropico3D` y `VariogramaCruzadoAnisotropico3D`, tiene dos propósitos. El primero es obtener un conjunto de coordenadas en un sistema de georeferenciación plano y el segundo, definir los parámetros de la región de búsqueda de información. En ambas clases se han implementado métodos para estimar el variograma direccional los cuales se han denominado `calcularVariogramaAnisotropico` y reciben dos parámetros con el propósito de definir la orientación del eje de la zona de búsqueda mostrada en la figura 5.1. Dichos parámetros corresponden a dos ángulos, que se han denominado acimut y dirección, α y β respectivamente en la figura 5.2. Los ángulos anteriores definen dos rotaciones de ejes coordenados. Inicialmente, α define el ángulo de rotación del eje X' alrededor del eje Z medido desde el eje Y , formando un nuevo sistema cartesiano de ejes coordenados (X', Y', Z) . Posteriormente, se realiza una nueva rotación, definida por el ángulo β que representa la rotación del eje X' alrededor del eje Y' medida desde el eje Z , formando un nuevo sistema cartesiano de ejes coordenados (X'', Y', Z') . La transformación de ejes implementada en este trabajo es similar a la presentada por Isaaks y Srivastava (1989). De esta forma, la zona de búsqueda tridimensional está dada por una región en el espacio definida por la rotación de la región del plano mostrada en la figura 5.1 alrededor del eje X'' , en la figura 5.2. Fuera de lo anterior, ambas clases tienen asociados métodos que realizan la lectura y la escritura de archivos de texto donde se consignan los resultados de las estimaciones de variograma, tal y como se ha comentado más arriba.

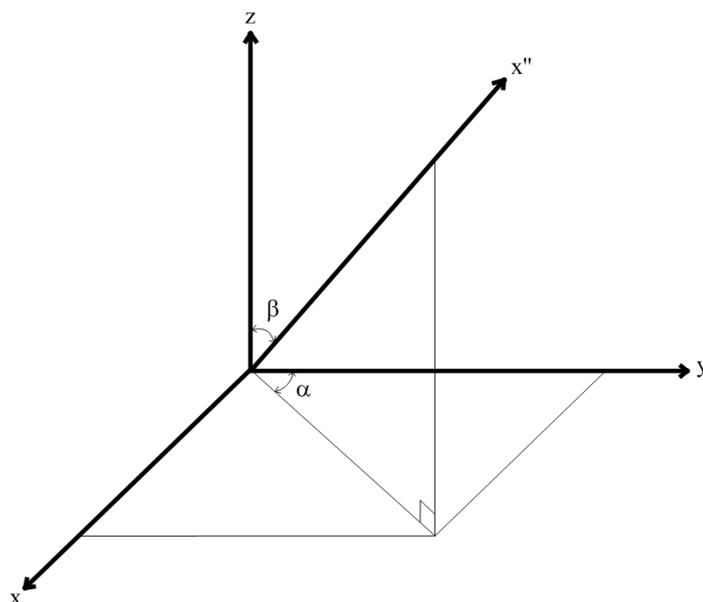


Figura 5.2. Esquema de los ángulos que definen la orientación del eje direccional del variograma

5.6. Variogramas superficiales

Un variograma superficial representa el conjunto de todos los variogramas direccionales que se pueden encontrar a partir de las observaciones disponibles. Su cálculo es muy útil para determinar, por inspección visual, las direcciones de máxima y mínima continuidad espacial. En un variograma superficial cada celda del mapa representa el valor del variograma experimental para una dirección y una distancia determinada (Isaaks y Srivastava, 1989; Goovaerts, 1997).

5.6.1. La clase MapaVariograma

Un objeto de la clase `MapaVariograma`, que se encuentra localizada en el paquete `JGeoST.variograma.objetos.experimental`, permite calcular un variograma superficial expresado como un objeto que puede pertenecer a las clases `MetaRaster` o `MetaRaster3D`, si se trata de análisis bidimensionales o tridimensionales, respectivamente. La clase tiene dos tipos de constructores, los cuales crean objetos de acuerdo a si el tipo de análisis de variograma a realizar es directo o cruzado. Así, en los constructores diseñados para un análisis de variograma directo no incluyen como parámetros las coordenadas de las localizaciones correspondientes a las realizaciones de la variable secundaria, los valores de dichas realizaciones y su sistema de georeferenciación. Por el contrario, para análisis de variogramas cruzados, los constructores correspondientes incluyen los parámetros previa-

mente mencionados. Así, los parámetros necesarios para construir un objeto de variograma superficial son:

- **Coordenadas de las localizaciones:** son las coordenadas de la ubicación de las realizaciones disponibles (en el caso del objeto de variograma cruzado también se definen las coordenadas para las realizaciones de la variable secundaria), las cuales están sujetas a un sistema de georeferenciación definido a priori, o simplemente son coordenadas referidas a un sistema ficticio local. Las coordenadas de las localizaciones pueden extraerse de un objeto de la clase `ListaObservaciones` utilizando los respectivos métodos de clase.
- **Valores de las realizaciones:** son los valores del atributo para las observaciones disponibles de la variable de análisis y/o la variable secundaria. Los valores de las realizaciones pueden extraerse de un objeto previamente construido perteneciente a la clase `ListaObservaciones` empleando los métodos asociados. Igualmente, este vector puede ser ingresado directamente como parámetro de clase y debe estar ordenado de forma que cada valor corresponda a las coordenadas de cada columna en la matriz de localizaciones.
- **Tolerancia angular:** es el ángulo máximo que define la zona de búsqueda para el variograma direccional, $\Delta\theta$ en la figura 5.1.
- **Ancho de banda:** es el ancho, en términos de distancia euclídea, que tiene la zona de búsqueda, b en la figura 5.1.
- **Georeferenciación de las localizaciones:** se trata de un objeto perteneciente a la clase `Georeferenciacion` que define el sistema de referenciación espacial para las coordenadas de las localizaciones asociadas con la información dura. En el caso de realizarse un análisis de variograma cruzado, debe definirse la georeferenciación de las localizaciones de la variable secundaria. Para ambos casos, si se trata de un sistema de coordenadas geográfico, las coordenadas se proyectan a un sistema de coordenadas plano, lo cual se lleva a cabo para corregir los errores de distancia que pueden presentarse al calcular distancias euclídeas sobre superficies curvas, como por ejemplo elipsoides de referencia para los datums geográficos. Si se está trabajando con un objeto de la clase `ListaObservaciones`, se extrae el sistema de georeferenciación de las coordenadas usando los métodos de clase asociados.
- **Sistema de georeferenciación destino:** se refiere a un objeto que pertenece a la clase `Georeferenciacion` que representa el sistema al cual se van a proyectar las coordenadas de las localizaciones para propósitos de cálculo de distancias euclídeas.

En la clase se han implementado cuatro métodos que realizan los cálculos del variograma de acuerdo a si se trata de un análisis directo, cruzado, bidimensional o tridimensional:

- **Método `getValoresMapaVariograma2D`:** calcula los valores del variograma superficial directo a partir de la información experimental y los ubica convenientemente en un vector que se usa como información para crear un objeto de la clase `MetaRaster`. Este objeto es el variograma superficial propiamente dicho, en el cual se representa un dominio de distancias bidimensionales.
- **Método `getValoresMapaVariograma3D`:** calcula los valores del variograma superficial directo a partir de la información experimental tridimensional y los ubica convenientemente en un vector que se usa como información para crear un objeto de la clase `MetaRaster3D`. Éste es el variograma propiamente dicho y representa un dominio de distancias tridimensional.
- **Método `getValoresMapaCruzadoVariograma2D`:** se trata de un método análogo a `getValoresMapaVariograma2D`, pero devuelve un objeto de la clase `MetaRaster` que representa el variograma superficial cruzado. El método debe usarse para objetos de la clase `MapaVariograma` que han sido creados ingresando como parámetros los objetos de la clase `ListaObservaciones` tanto para la variable principal como para la variable secundaria.
- **Método `getValoresMapaCruzadoVariograma3D`:** se trata de un método análogo a `getValoresMapaVariograma3D`, pero devuelve un objeto de la clase `MetaRaster3D` que representa el variograma cruzado. De la misma forma que el método `getValoresMapaCruzadoVariograma2D`, debe usarse para objetos de la clase `MapaVariograma` que han sido creados ingresando como parámetros los objetos de la clase `ListaObservaciones`, tanto para la variable principal como para la variable secundaria.

Las salidas de todos los métodos son objetos de manipulación de entidades raster, pertenecientes a las clases `MetaRaster` y `MetaRaster3D`, por lo cual tienen asociados los métodos incluidos en ellas.

5.7. Ajuste de modelos a los variogramas experimentales

Una vez calculados los variogramas experimentales es necesario ajustar a ellos una función que nos permita conocer el variograma para todo vector separación. Sin embargo, no se puede emplear cualquier función ya que es necesario garantizar que las matrices que se construyan en cualquiera de las variantes de krigeado sean definidas positivas (Armstrong y Jabin, 1981; Isaaks y Srivastava, 1989; Goovaerts, 1997). Para conseguir lo anterior se deben utilizar modelos que garanticen esa propiedad. Por ello, en esta aplicación se han incluido algunos de los modelos básicos permitidos de variograma más frecuentemente empleados:

- **Modelo efecto pepita:**

$$g(\mathbf{h}) = \begin{cases} 0 & \text{Si } |\mathbf{h}| = 0 \\ b & \text{Si } |\mathbf{h}| > 0 \end{cases} \quad (5.5)$$

- **Modelo esférico:**

$$Esf(\mathbf{h}) = \begin{cases} 1.5b\frac{|\mathbf{h}|}{a} - 0.5b\left(\frac{|\mathbf{h}|}{a}\right)^3 & \text{Si } |\mathbf{h}| \leq a \\ 1 & \text{Si } |\mathbf{h}| > a \end{cases} \quad (5.6)$$

- **Modelo lineal acotado:**

$$Lin(\mathbf{h}) = \begin{cases} b\frac{|\mathbf{h}|}{a} & \text{Si } |\mathbf{h}| \leq a \\ b & \text{Si } |\mathbf{h}| > a \end{cases} \quad (5.7)$$

- **Modelo exponencial con un rango práctico a :**

$$Exp(\mathbf{h}) = b \left[1 - \exp\left(-\frac{3|\mathbf{h}|}{a}\right) \right] \quad (5.8)$$

- **Modelo gaussiano con un rango práctico a :**

$$Gau(\mathbf{h}) = b \left[1 - \exp\left(-\frac{3|\mathbf{h}|^2}{a^2}\right) \right] \quad (5.9)$$

En las anteriores ecuaciones, a es el alcance, es decir, la distancia hasta la cual existe correlación entre valores de la variable; $|\mathbf{h}|$ es el vector separación; y b es la meseta, es decir, el máximo valor del variograma. Para evaluar cual es el mejor modelo que represente las características del variograma experimental existen muchos métodos de ajuste que pueden aplicarse, por ejemplo la regresión mediante mínimos cuadrados ordinarios, mínimos cuadrados generalizados, mínimos cuadrados ponderados y el análisis visual. En ESTELLUS se han implementado algoritmos para el ajuste de variogramas mediante mínimos cuadrados ordinarios y mínimos cuadrados ponderados y se ha implementado un mecanismo interactivo para realizar un análisis de ajuste visual, el cual se presentará con detalle en el capítulo 7.

5.7.1. Manejo computacional de los modelos de variograma

La manipulación de los diferentes modelos de variograma se realiza mediante el uso de una interface Java denominada `RegressionFunction`, que encuentra dentro del paquete `JGeoST.variograma.modelo`. Las interfaces Java son expresiones puras de diseño. Éstas son auténticas conceptualizaciones no implementadas que sirven de guía para definir un determinado concepto referido a una clase Java y lo que debe éste hacer, pero sin desarrollar un mecanismo de solución explícito. Se trata de declarar métodos abstractos y

constantes que posteriormente puedan ser implementados de diferentes maneras según las necesidades de un programa. De este modo, la interface `RegressionFunction` implementa los siguiente métodos abstractos:

- **function:** es un método que recibe los parámetros del modelo y un vector de distancias, para las cuales devuelve los valores correspondientes del variograma asociado. Este método se usa para ser incluido en el procedimiento de ajuste automático al variograma experimental, como se muestra en la siguiente sección.
- **evaluate:** es un método que recibe los parámetros del modelo y evalúa el variograma asociado para una distancia dada. Este método se usa para la creación de superposiciones lineales de modelos de variograma anisotrópicos.
- **covariance:** este método evalúa la función de covarianza correspondiente al modelo de variograma. Este tipo de funciones se incluyen para evaluar superposiciones anisotrópicas de estructuras imbricadas tridimensionales en los procedimientos de estimación, para algunos tipos de krigeados.

Aparte de de la interface `RegressionFunction`, se tienen las clases `Pepita`, `Lineal`, `Gaussiano`, `Exponencial` y `Esferico`, las cuales implementan los métodos previamente comentados. En cada método correspondiente a cada clase se han adaptado las expresiones matemáticas para la evaluación de los modelos de variograma, como se presentó anteriormente en las ecuaciones 5.5, 5.6, 5.7, 5.8 y 5.9.

5.7.2. Ajuste mediante mínimos cuadrados ordinarios

El método de ajuste mediante los mínimos cuadrados ordinarios se basa en la minimización de una función objetivo que involucra la diferencia cuadrática entre los valores del variograma experimental y los correspondientes al modelo ajustado, como se muestra en la siguiente expresión:

$$MCO(\tau, \gamma) = \sum_{j=1}^k [\hat{\gamma}_j(\mathbf{h}_j) - \gamma(\mathbf{h}_j, \tau)]^2 \quad (5.10)$$

donde τ es el vector de parámetros del modelo, \mathbf{h}_j es la distancia a la cual se desea evaluar el variograma, γ es el modelo de variograma que se desea ajustar y $\hat{\gamma}_j(\mathbf{h}_j)$ es el valor del variograma experimental para la distancia \mathbf{h}_j .

5.7.3. Ajuste mediante mínimos cuadrados ponderados

El método de los mínimos cuadrados ponderados se basa en la minimización de una función objetivo que involucra la diferencia cuadrática entre los valores del variograma

experimental y los correspondientes al modelo multiplicados por un factor de ponderación, de la siguiente manera:

$$MCP(\tau, \gamma) = \sum_{j=1}^k \frac{[\hat{\gamma}_j(\mathbf{h}_j) - \gamma(\mathbf{h}_j, \tau)]^2}{\text{Var}[\hat{\gamma}_j(\mathbf{h}_j)]} \quad (5.11)$$

donde τ es el vector de parámetros del modelo, \mathbf{h}_j es la distancia a la cual se desea evaluar el variograma, γ es el modelo que se desea ajustar, $\hat{\gamma}_j(\mathbf{h}_j)$ es el valor del variograma experimental para la distancia \mathbf{h}_j y $\text{Var}[\hat{\gamma}_j(\mathbf{h}_j)]$ es la varianza de estimación asociada con el variograma experimental. El valor anterior, en general, es desconocido pero puede determinarse de diferentes maneras. Un criterio que ha sido ampliamente utilizado es el propuesto por Cressie (1985), quien afirma que el valor de dicha varianza se aproxima a:

$$\text{Var}[\hat{\gamma}_j(\mathbf{h}_j)] = \frac{[\hat{\gamma}(\mathbf{h}_j)]^2}{n_j} \quad (5.12)$$

donde n_j es el número de pares de realizaciones de la variable analizadas que se encuentran a una distancia \mathbf{h}_j . Años después, Zhang et al. (1995) propusieron una modificación de la ecuación 5.12, principalmente para evitar que los factores de ponderación sean función de los parámetros a determinar en el ajuste, lo cual ocasiona una convergencia más lenta del algoritmo, sugiriendo que:

$$\text{Var}[\hat{\gamma}_j(\mathbf{h}_j)] = \frac{[\mathbf{h}_j]^2}{n_j} \quad (5.13)$$

Ambas expresiones de ponderación se incluyen en las opciones de la herramienta de ajuste automático de variograma, como se presenta en el capítulo 7.

5.7.4. El algoritmo de optimización *Nelder-Mead*

El ajuste de los modelos a los variogramas experimentales por mínimos cuadrados y mínimos cuadrados ponderados se hace utilizando el algoritmo de optimización *Nelder-Mead* o *simplex*, publicado originalmente por Nelder y Mead (1965). Se trata de una de las técnicas más utilizados para la optimización sin restricciones que no requieren el uso de expresiones matemáticas explícitas para las derivadas de la función a optimizar. El algoritmo fue diseñado para resolver el problema clásico de minimizar un función no lineal $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y sólo usa valores de la función en algunos puntos de \mathbb{R}^n . Tampoco usa o aproxima gradientes de la función en dichos puntos, por lo cual ha sido catalogado como un método de optimización perteneciente a los métodos de búsqueda directa (Press et al., 1988).

Como otros métodos de búsqueda directa, el método *Nelder-Mead* está basado en un *simplex* o región de búsqueda. Un *simplex* S en \mathbb{R}^n se define como espacio convexo de $n + 1$ vértices $x_0, \dots, x_n \in \mathbb{R}^n$. Por ejemplo, un *simplex* en \mathbb{R}^2 es un triángulo y en \mathbb{R}^3 es un tetraedro.

El método efectúa una serie de transformaciones del *simplex* S , de forma que el valor de la función objetivo decrece en sus vértices. En cada paso, la transformación se determina calculando uno o más puntos de prueba y los valores correspondientes de las funciones objetivo y comparándolos con los existentes en los vértices del *simplex*. Los detalles operativos e implementaciones computacionales del método pueden encontrarse en el texto de Press et al. (1988).

5.7.5. La clase AjustarMinimosCuadradosNelderMead

En la librería de ESTELLUS se ha creado una clase que permite el ajuste de los modelos de variograma mediante mínimos cuadrados y mínimos cuadrados ponderados utilizando el algoritmo de optimización *Nelder-Mead*. La clase creada, denominada como `AjustarMinimosCuadradosNelderMead`, se encuentra dentro del paquete `JGeosST.variograma.objetos.modelo`. Es una adaptación del método `neldermead` existente en la clase `Regression` de la librería del profesor Flanagan (2008). Los parámetros para construir un objeto de la clase de ajuste por mínimos cuadrados ordinarios o ponderados son:

- **Modelo a ajustar:** es un objeto que implementa la interface `RegressionFunction` y representa el tipo de modelo que se desea ajusta al variograma experimental.
- **Booleano de ponderación:** es una variable booleana que indica si el ajuste se realizará utilizando mínimos cuadrados ordinarios o ponderados.
- **Datos de x:** es un vector donde se almacenan los valores de las distancias para las cuales se ha evaluado el variograma experimental.
- **Datos de y:** es un vector donde se almacenan los valores del variograma experimental.
- **Ponderaciones:** son los valores de las ponderaciones que serán incluidas al inicio del algoritmo, por ejemplo, el criterio de Cressie (1985).

La clase tiene asociada el método `neldermead`, adaptado para permitir la inclusión de los modelos permitidos de variograma. En la clase se han creado diferentes versiones del método de acuerdo con los parámetros requeridos para la ejecución del método numérico; así, el algoritmo computacional que implementa la versión más general del método de optimización *Nelder-Mead* requiere los siguientes parámetros:

- **Parámetros iniciales:** se trata de un conjunto de valores iniciales de los parámetros de los modelos de variograma, rangos y mesetas, a partir de los cuales empieza la ejecución del método de búsqueda.
- **Tamaño inicial de paso de cada parámetro:** es el incremento inicial al cual será sometido cada parámetro en cada iteración asociada con la deformación del *simplex* de búsqueda.

- **Tolerancia:** es el error mínimo relativo para el cambio de la función objetivo en iteraciones sucesivas, sea mínimos cuadrados o mínimos cuadrados ponderados, para el cual el método termina la búsqueda.
- **Número de iteraciones:** es el número máximo de iteraciones que puede llevar a cabo el algoritmo de búsqueda. Al llegar a este número de iteraciones la búsqueda del mínimo cuadrático termina a pesar de que no se haya alcanzado la tolerancia especificada.

Finalmente, el método devuelve los valores de los parámetro de las funciones permitidas de variograma que minimizan la función objetivo para un determinado variograma experimental.

5.8. Modelos anisotrópicos de variograma

Cuando la estructura espacial de un fenómeno depende de la dirección se dice que éste es anisótropo. En la práctica, la presencia de anisotropía se puede detectar analizando los variogramas direccionales. Se puede hablar de tres tipos de anisotropía: geométrica, zonal y una combinación de ambas. Cuando los variogramas en diferentes direcciones presentan la misma forma y meseta, pero diferente rango, se habla de anisotropía geométrica. Si existe una diferencia en las mesetas con la dirección, se dice que la anisotropía es zonal. La presencia de anisotropía combinada se detecta al observar una variación simultánea de rangos y mesetas con respecto a la dirección (Goovaerts, 1997).

Un modelo anisotrópico de continuidad espacial parte del conocimiento de los modelos de variograma ajustados en las direcciones correspondientes a los ejes de máxima y mínima anisotropía, los cuales definen un sistema de ejes coordenados que en general no coincide con el sistema de georeferenciación empleado para la localización de los datos. Según Isaaks y Srivastava (1989) y Goovaerts (1997), el siguiente paso consiste en encontrar un modelo que describa como cambia el variograma con respecto a la dirección analizada. Los dos textos antes citados, muestran expresiones matemáticas para el tratamiento de los diferentes tipos de anisotropía, a pesar de ello, en ESTELLUS sólo se han implementado ecuaciones de transformación para anisotropía geométrica por dos razones fundamentales:

1. Puede demostrarse que la anisotropía zonal corresponde a un modelo anisotrópico geométrico cuyo rango mayor tiende a infinito. Así, para la modelación de la anisotropía zonal basta con asignar un valor muy elevado del rango en la dirección de máxima continuidad.
2. La anisotropía combinada puede construirse mediante la superposición de modelos anisotrópicos geométricos y zonales.

El algoritmo para la construcción de un modelo anisotrópico tridimensional implementado en ESTELLUS se basa en los conceptos propuestos por Isaaks y Srivastava (1989) y consta de los siguiente pasos:

- Ajuste de modelos de variograma en las direcciones de anisotropía.
- Definición de la transformación de rango unitario estandarizado, es decir, se plantea una expresión que reduce los variogramas direccionales a un modelo común de rango unitario. En tres dimensiones debe emplearse un modelo de variograma $\gamma(\mathbf{h}) = \gamma(h_x, h_y, h_z) = \gamma(h_1)$, con la distancia transformada así:

$$h_1 = \sqrt{\left\{\frac{h_x}{a_x}\right\}^2 + \left\{\frac{h_y}{a_y}\right\}^2 + \left\{\frac{h_z}{a_z}\right\}^2} \quad (5.14)$$

- Finalmente, se realiza la rotación de los ejes coordenados de forma que éstos coincidan con los ejes de anisotropía. Para lograr lo anterior se definen tres ángulos, como se presenta en las figuras 5.3, 5.4 y 5.5:
 1. α es el ángulo que define la rotación de los ejes X e Y alrededor del eje Z (denominado acimut), para obtener un nuevo sistema de ejes coordenados (X', Y', Z) .
 2. β es el ángulo que define la rotación de los ejes X' y Z alrededor del eje Y' para obtener un nuevo sistema de ejes coordenados (X'', Y', Z') .
 3. θ es el ángulo que define la rotación de los ejes Y' y Z' alrededor del eje X'' para obtener un nuevo sistema de ejes coordenados (X'', Y'', Z'') .

La distancia transformada por rotación se calcula como:

$$\mathbf{h}' = \mathbf{R}_3 \mathbf{R}_2 \mathbf{R}_1 \mathbf{h} \quad (5.15)$$

donde \mathbf{R}_i , con $i = 1, 2, 3$, son matrices de transformación, definidas como (Remy, 2004):

$$\mathbf{R}_1 = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5.16)$$

$$\mathbf{R}_2 = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix} \quad (5.17)$$

$$\mathbf{R}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{bmatrix} \quad (5.18)$$

A partir de las expresiones 5.14, 5.16, 5.17 y 5.18 se puede evaluar el valor del variograma para cualquier distancia en la dirección de tres nuevos ejes coordenados cartesianos en el espacio tridimensional. Dichos ejes corresponden entonces a las direcciones de anisotropía encontradas en el análisis experimental de los variogramas.

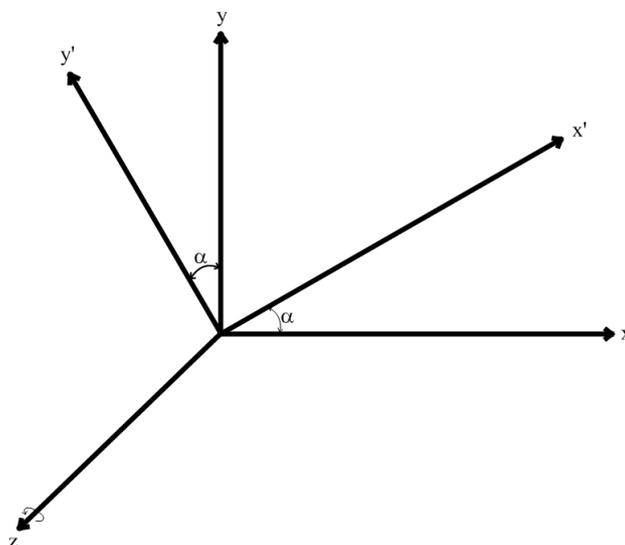


Figura 5.3. Esquema de la primer rotación de los ejes coordenados para lograr que éstos coincidan con los ejes de anisotropía. Se define el ángulo α , denominado acimut, para rotar los ejes X y Y alrededor del eje Z . De esta forma se obtiene un nuevo sistema coordenado (X', Y', Z)

5.8.1. La clase EvaluarVariograma

Para evaluar el valor del variograma correspondiente a cualquier distancia euclídea en el espacio tridimensional, se han implementado las clases `EvaluarVariograma`, `EvaluarAnisotropia` y `SuperposicionAnisotropia`. La primera evalúa el variograma sin tener en cuenta la posible anisotropía. La clase `EvaluarVariograma` está contenida en el paquete `JGeoST.variograma.objetos.modelo`. El constructor de dicha clase recibe los siguientes parámetros:

- **Parámetros del modelo:** es un vector en el cual se almacenan los parámetros del modelo en el orden: `[meseta, rango]`.
- **Tipo de modelo:** se trata del nombre de archivo de clase que contiene el código del modelo correspondiente. Por ejemplo, el archivo `*.java` correspondiente al modelo esférico se llama `Esferico.java` y se localiza en el paquete `JGeoST.variograma.modelo`. El constructor localiza la ubicación del paquete y construye virtualmente un objeto a partir del nombre del archivo que lo contiene. Este tipo de conceptualización computacional diseñada para este constructor hace posible que el programador pueda seguir añadiendo diferentes tipos de modelos fuera de los existente y la librería continúe en crecimiento permanente.

La clase dispone del método `getVariograma`, cuyo único parámetro es una distancia, para la cual el método evalúa el variograma correspondiente.

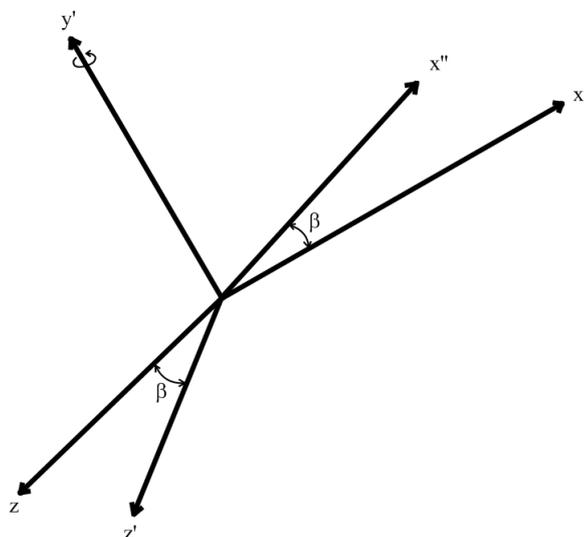


Figura 5.4. Esquema de la segunda rotación de los ejes coordenados para lograr que éstos coincidan con los ejes de anisotropía. Se define el ángulo β , denominado dirección, para rotar los ejes X' y Z' alrededor del eje Y' . De esta forma se obtiene un nuevo sistema coordenado (X'', Y', Z')

5.8.2. La clase `EvaluarAnisotropia`

El objeto que permite llevar a cabo la evaluación del variograma anisotrópico se construye en la clase `EvaluarAnisotropia`, que se encuentra contenida en el interior del paquete `JGeoST.variograma.objetos.modelo`. En ella se han implementado los métodos que calculan los valores del variograma para una orientación dada a partir de los modelos ajustados en las direcciones de anisotropía. Como se comentó anteriormente, esta clase solamente evalúa los efectos de la anisotropía geométrica, con lo cual las mesetas correspondientes a los objetos de `EvaluarVariograma` deben ser iguales. Si se desea modelar anisotropía zonal, debe asignarse un valor al rango en la dirección de máxima continuidad mucho mayor que los asociados a las restantes direcciones de anisotropía y asignarlo al parámetro correspondiente en el objeto de la clase `EvaluarVariograma`. El método básicamente aplica las ecuaciones de transformación 5.14, 5.16, 5.17 y 5.18 para evaluar el variograma para una rotación y una distancia dada. Los parámetros requeridos para la construcción de un objeto de variograma anisotrópico perteneciente a la clase `EvaluarAnisotropia` son:

- **Modelos ajustados previamente en las direcciones de anisotropía:** corresponden a objetos de la clase `EvaluarVariograma` que permiten realizar el cálculo de los valores del modelo para una distancia dada.

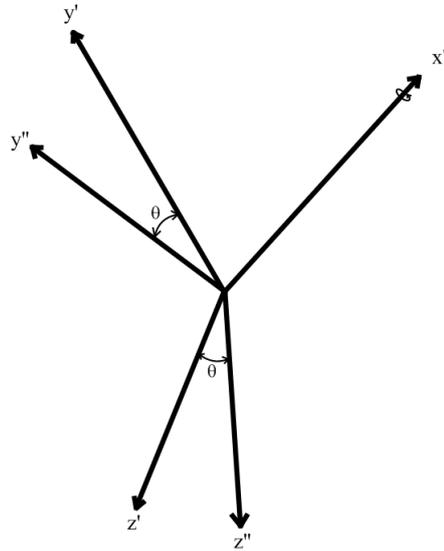


Figura 5.5. Esquema de la tercera rotación de los ejes coordenados para lograr que éstos coincidan con los ejes de anisotropía. Se define el ángulo θ , denominado inmersión, para rotar los ejes Y' y Z' alrededor del eje X'' . De esta forma se obtiene un nuevo sistema coordenado (X'', Y'', Z'')

- **Acimut (α):** es el ángulo de rotación de los ejes X y Y alrededor del eje Z . El dominio de esta ángulo es $[0^\circ, 180^\circ]$, su valor debe ser introducido en grados.
- **Dirección (β):** es el ángulo de rotación de los ejes X' y Z alrededor del eje Y' . El dominio de esta ángulo es $[0^\circ, 180^\circ]$, su valor debe ser introducido en grados.
- **Inmersión (θ):** es el ángulo de rotación de los eje Y' y Z' alrededor del eje X'' . El dominio de esta ángulo es $[0^\circ, 180^\circ]$, su valor debe ser introducido en grados.

Todas las rotaciones se llevan a cabo siguiendo la regla de la mano derecha. El método de clase que contiene explícitamente los pasos enumerados al inicio de esta sección y lleva a cabo el cálculo del variograma anisotrópico para una distancia vectorial (h_x, h_y, h_z) dada se llama `getSemivariogramaAnisotropico`. Los parámetro requeridos por dicho método son precisamente los componentes de la distancia vectorial a la cual se desea estimar el variograma, para la orientación determinada previamente.

Operativamente, la clase permite únicamente la evaluación del modelo anisotrópico para un solo tipo de modelo de variograma, es decir, dado que se disponga de una estructura imbricada, un objeto de esta clase representaría sólo uno de los componentes de la superposición lineal. Es posible aplicar el principio de superposición sobre objetos de esta clase para obtener estructuras imbricadas. A continuación se presenta una clase que hace posible la superposición de estructuras anisotrópicas de variograma.

5.8.3. La clase SuperposicionAnisotropia

La clase `SuperposicionAnisotropia`, que se encuentra localizada en el interior del paquete `JGeoST.variograma.objetos.modelo`, permite la evaluación del valor del variograma correspondiente a estructuras imbricadas anisotrópicas mediante la aplicación del principio de superposición sobre estructuras anisotrópicas básicas que intervienen en ella. La construcción de un objeto para la modelación de estructuras imbricadas anisotrópicas recibe los siguientes parámetros:

- **Estructuras anisotrópicas involucradas:** es un vector que contiene las estructuras anisotrópicas que se van a superponer en la estructura imbricada. Así, las componentes de dicho vector son objetos de la clase `EvaluarAnisotropia`.
- **Efecto pepita:** Es el valor de la meseta de la estructura isotrópica del efecto pepita. Si la modelación no considera dicha estructura debe asignarse un valor de cero.

La evaluación del valor del variograma y/o covarianza se realiza mediante los métodos `getVariograma` y `getCovarianza`, los cuales reciben el vector de distancia, (h_x, h_y, h_z) como parámetro de cálculo. El método opera evaluando el valor del variograma o covarianza para cada estructura independientemente y superpone linealmente dichos valores, adicionando la influencia del efecto pepita, si se ha considerado en la modelación.

5.9. Comentarios finales

En este capítulo se han descrito con detalle las clases que hacen posible el análisis estructural de la(s) variable(s) a utilizar en el procedimiento de estimación. Así los pasos que se deberían seguir para realizar de forma adecuada el análisis estructural son:

- Calcular el variograma omnidireccional.
- Calcular los variogramas experimentales para diferentes direcciones y el variograma superficial, con el fin de detectar los ejes de anisotropía.
- Halladas las direcciones principales de anisotropía, calcular los variogramas experimentales direccionales correspondientes.
- Ajustar un modelo a las direcciones principales de anisotropía.
- En caso de que fuera necesario, en base a los variogramas ajustados, construir un modelo de correogionalización.

En el capítulo 7 se muestran las herramientas interactivas diseñadas para llevar a cabo el análisis estructural mediante ESTELLUS.

6

Estimación espacial

6.1. Introducción

En este capítulo se presentan las bases conceptuales de los algoritmos de estimación espacial incluidos en ESTELLUS, que son los siguientes: krigeado simple, krigeado ordinario, krigeado con deriva externa, cokrigeado ordinario, cokrigeado ordinario estandarizado y cokrigeado colocalizado markoviano. Dichos algoritmos se presentan según sean para la interpolación de una variable a partir de información de la misma o de una variable incorporando información externa. Se han programado también un par de técnicas de interpolación determinista como lo son el método del vecino cercano y el método del inverso de la distancia elevado a un exponente.

6.2. Algoritmos de estimación determinista

Los métodos de estimación deterministas ajustan una superficie matemática, de forma previamente determinada, a un conjunto de puntos con valores conocidos para una variable espacial. Así pues, los modelos tradicionales de interpolación deterministas no proporcionan herramientas para cuantificar la incertidumbre o para modelar el patrón de variabilidad espacial de una variable.

6.2.1. Método del vecino cercano

El método del vecino cercano asigna a cada celda del campo raster el valor de la observación más cercana mediante la construcción de los *Polígonos de Thiessen*, también

conocidos como *Polígonos de Voronoi* o *Teselación de Dirichlet*, los cuales son un mero ejercicio geométrico que permite construir una partición del plano interpolado. Es uno de los métodos de interpolación más simples, basado en la distancia euclídea, siendo especialmente apropiado cuando los datos son cualitativos. Se trata de unir los puntos interpolados entre sí y trazar las mediatrices de los segmento de unión. Las intersecciones de estas mediatrices determinan una serie de polígonos en un espacio bidimensional que permite asignar a cada punto un área de influencia. El valor de la variable dentro de cada polígono es el de la estación contenida (Burrough y McDonnell, 1998).

6.2.2. Método del inverso de la distancia elevado a un exponente

El método de interpolación de medias móviles ponderadas por la distancia es ampliamente usado en la modelación de superficies. Se basa en la idea intuitiva, conocida como la *Ley de Tobler* de la geografía (Burrough y McDonnell, 1998), de que las observaciones más cercanas deben tener más peso en la determinación del valor interpolado en un determinado punto. Se trata de un método exacto y local que estima un valor de la variable z^* en un punto x_0 a partir de la siguiente expresión:

$$z^*(x_0) = \frac{\sum_{i=0}^n w(d_i)z(x_i)}{\sum_{i=0}^n w(d_i)} \quad (6.1)$$

donde $w(d_i)$ denota la función peso de los n puntos muestrales que intervienen en el cálculo y d_i representa la distancia de cada punto a x_0 . Las funciones de ponderación adoptan la forma general $w = d^{-u}$ y los valores más usuales del exponente son $u = 1$ y $u = 2$ (Lam, 1983), aunque algunos autores establecen que el exponente 2 suele ofrecer unos resultados empíricos más satisfactorios, además de requerir un menor esfuerzo computacional (Declercq, 1996). Obviamente el uso de exponentes elevados implica dar menos peso relativo a puntos más alejados de x_0 y viceversa. Probablemente el mayor problema que presenta este método es que los valores interpolados son medias ponderadas que siempre toman valores entre el máximo y el mínimo de los puntos muestrales.

6.2.3. Los objetos computacionales para la estimación determinista

La estimación mediante los dos métodos deterministas comentados se ha implementado en cuatro clases de Java, denominadas `Vecino2D`, `Vecino3D`, `Inverso2D` e `Inverso3D`, que se encuentran contenidas dentro del paquete `JGeoST.estimadores.univariados.objetos`. La función del constructor de las clases es recibir las coordenadas y los valores de los datos disponibles, mediante un objeto de la clase `ListaObservaciones` y hacer las conversiones necesarias para calcular correctamente las distancias euclídeas. Los parámetros para los diferentes constructores de las clases implementadas son:

- **Coordenadas de los localizaciones:** es una matriz en la que se almacenan las coordenadas de las localizaciones donde se dispone de valores de la variable analiza-

da. Pueden extraerse de un objeto de la clase `ListaObservaciones` utilizando los métodos de clase correspondientes.

- **Valores de las observaciones:** vector en donde se consignan los valores de las realizaciones de la variable analizada. Pueden extraerse de un objeto previamente construido perteneciente a la clase `ListaObservaciones` empleando los métodos asociados. Igualmente, este vector puede ser ingresado directamente como parámetro de clase y debe estar ordenado de forma que cada valor corresponda a las coordenadas de cada columna en la matriz de localizaciones.
- **Georeferenciación de las coordenadas:** es un objeto construido mediante la clase `Georeferenciacion` que indica cual es el sistema de coordenadas usado para ubicar las coordenadas de las localizaciones. Si se están usando objetos de la clase `ListaObservaciones`, el sistema de georeferenciación puede extraerse empleando los métodos de dicha clase.
- **Georeferenciación destino:** es un objeto de la clase `Georeferenciacion` que se usa en el caso de que se estén usando coordenadas geográficas para proyectar las coordenadas a un sistema plano y calcular las distancias.
- **Exponente:** En los casos de los constructores para el método del inverso de la distancia a un exponente, debe asignarse su valor mediante este parámetro.

El método encargado de la estimación del campo raster se ha llamado `estimarMalla`, que tiene como parámetros:

- **Georeferenciación original y de destino:** son objetos construidos usando la clase `Georeferenciacion` que indican el sistema de georeferenciación de las coordenadas asociadas a las localizaciones con información y el sistema de georeferenciación que se desea para el campo interpolado resultante de la aplicación del método.
- **Metadatos del raster de salida:** son las características que definen el dominio espacial del campo raster de salida del método: las coordenadas de la esquina inferior izquierda, el tamaño de las celdas en cada una de las tres direcciones coordenadas y el número de celdas.

6.3. Algoritmos de estimación estocástica de una variable

Los algoritmos de estimación geoestadística de una variable considerando la información que de esa variable se encuentra disponible se engloban bajo el nombre genérico de kriging. La idea básica del kriging es estimar el valor desconocido de un atributo z en un punto donde no se disponen mediciones y que está localizado en \mathbf{u} , como una combinación lineal de n valores conocidos de dicho atributo, cuyas coordenadas son \mathbf{u}_α , con $\alpha = 1, \dots, n$. Para representar el parámetro o atributo de interés, z se adopta un modelo

estocástico que consiste en admitir que los valores que puede tomar z dentro del área de estudio vienen dados por una función aleatoria $Z(\mathbf{u})$. La forma general del estimador por krigado viene expresada a través de la siguiente ecuación:

$$Z^*(\mathbf{u}) - m(\mathbf{u}) = \sum_{\alpha=1}^n \lambda_{\alpha}(\mathbf{u}) [Z(\mathbf{u}_{\alpha}) - m(\mathbf{u}_{\alpha})] \quad (6.2)$$

donde el símbolo * denota que se trata de un estimador; $m(\mathbf{u})$ y $m(\mathbf{u}_{\alpha})$ son los valores esperados de las variables aleatorias $Z(\mathbf{u})$ y $Z(\mathbf{u}_{\alpha})$ respectivamente, y $\lambda(\mathbf{u}_{\alpha})$, $\alpha = 1, \dots, n$ son los coeficientes de ponderación asignados a los datos de la variable que intervienen en la estimación. La representación estocástica del atributo $Z(\mathbf{u})$, permite definir el error en la estimación realizada como:

$$R(\mathbf{u}) = Z^*(\mathbf{u}) - Z(\mathbf{u}) \quad (6.3)$$

donde $R(\mathbf{u})$ es el error cometido en la estimación del atributo analizado en la localización \mathbf{u} ; $Z^*(\mathbf{u})$ y $Z(\mathbf{u})$ son las variables aleatorias que representan el valor estimado y el valor verdadero, respectivamente. Los pesos del krigado se determinan imponiendo que:

$$\text{Min } \{\sigma_R^2(\mathbf{u})\} = \text{Min } \{\text{Var}[Z^*(\mathbf{u}) - Z(\mathbf{u}_0)]\} \quad (6.4)$$

con la restricción:

$$E[R(\mathbf{u})] = E[Z^*(\mathbf{u}) - Z(\mathbf{u})] = 0 \quad (6.5)$$

es decir que la varianza de los errores cometidos en un conjunto de estimaciones debe ser mínima y el estimador insesgado.

Todos los estimadores por krigado son una variante de la expresión general dada en la ecuación 6.2, en la cual los coeficientes de ponderación se obtienen resolviendo el sistema de ecuaciones lineales que resulta a partir de imponer las condiciones mostradas en las ecuaciones 6.4 y 6.5. A continuación se describe brevemente las principales variantes del krigado que han sido implementadas en ESTELLUS. Para encontrar una descripción exhaustiva de cada algoritmo, el lector interesado puede consultar cualquier texto clásico de Geoestadística (Isaaks y Srivastava, 1989; Goovaerts, 1997; Burrough y McDonnell, 1998; Deutsch y Journel, 1998; Chilès y Delfiner, 1999).

6.3.1. Kriging simple

El estimador por krigado simple (KS) tiene la siguiente expresión:

$$Z_{KS}^*(\mathbf{u}) - m = \sum_{\alpha=1}^n \lambda_{\alpha}^{KS}(\mathbf{u}) [z(\mathbf{u}_{\alpha}) - m] \quad (6.6)$$

donde m , que es igual al valor esperado de la variable aleatoria $Z(\mathbf{u})$, se considera conocido y constante en todo el dominio de la estimación. En la práctica, éste se asimila con el valor

promedio de los datos disponibles. Los coeficientes de ponderación $\lambda_\alpha^{KS}(\mathbf{u})$ se obtienen al exigir que el estimador sea insesgado y que la varianza de los errores sea mínima, lo cual conduce al siguiente sistema lineal de ecuaciones lineales, denominado sistema de *KS*:

$$\sum_{\beta=1}^n \lambda_\beta^{KS}(\mathbf{u})C(\mathbf{u}_\alpha - \mathbf{u}_\beta) = C(\mathbf{u}_\alpha - \mathbf{u}), \text{ con } \alpha = 1, 2, \dots, n(\mathbf{u}) \quad (6.7)$$

donde $C(\mathbf{h})$ es la covarianza para un vector separación \mathbf{h} . En este sistema, a la izquierda de la igualdad están las covarianzas entre los datos y a la derecha entre los datos y el punto a estimar. De acuerdo con la varianza de los errores, la incertidumbre de la interpolación por *KS* resulta:

$$\sigma_{KS}^2(\mathbf{u}) = \sigma^2 - \sum_{\beta=1}^n \lambda_\beta^{KS}(\mathbf{u})C(\mathbf{u}_\beta - \mathbf{u}) \quad (6.8)$$

donde σ^2 es la varianza correspondiente a los datos, o sea $C(\mathbf{0})$.

6.3.2. Kriging ordinario

El estimador por krigeado ordinario (*KO*) viene dado por la siguiente expresión:

$$z_{KO}^*(\mathbf{u}) = \sum_{\alpha=1}^n \lambda_\alpha^{KO}(\mathbf{u})z(\mathbf{u}_\alpha) \quad (6.9)$$

Los criterios exigidos al krigeado conducen al sistema de ecuaciones lineales del *KO*:

$$\begin{cases} \sum_{\beta=1}^n \lambda_\beta^{KO}(\mathbf{u})C(\mathbf{u}_\alpha - \mathbf{u}_\beta) + \mu^{KO} = C(\mathbf{u}_\alpha - \mathbf{u}), \text{ con } \alpha = 1, 2, \dots, n(\mathbf{u}) \\ \sum_{\beta=1}^n \lambda_\beta^{KO}(\mathbf{u}) = 1 \end{cases} \quad (6.10)$$

donde se puede observar que en este caso los pesos deben sumar uno, condición necesaria para obtener un estimador insesgado. El valor μ^{KO} es el parámetro de Lagrange introducido en la minimización del error sujeto a la condición de que los pesos sumen uno. La expresión para la varianza de la estimación por *KO* es la siguiente:

$$\sigma_{KO}^2(\mathbf{u}) = \sum_{\beta=1}^n \lambda_\beta^{KO}(\mathbf{u})C(\mathbf{u}_\beta - \mathbf{u}) - \mu^{KO} \quad (6.11)$$

6.3.3. Los objetos computacionales para la estimación por krigeado simple y ordinario

Para realizar la estimación de una variable en dominios bi y tridimensionales se han implementado cuatro clases *Java*, las cuales han sido nombradas como *KrigingSim2D*,

KrigingSim3D, KrigingOrd2D y KrigingOrd3D, y pertenecen al paquete `JGeoST.estimadores.univariados.objetos`. Todas las clases disponen de un constructor cuya función básica es recibir las coordenadas de las localizaciones y los valores correspondientes a las realizaciones de la variable analizada, extraídos desde un objeto de la clase `ListaObservaciones`. Posteriormente, y dado el caso de que la georeferenciación de las coordenadas sea geográfica (*longitud, latitud, prof.*), el constructor aplica la transformación conveniente para realizar el cálculo de vectores separación en base a distancias euclídeas, lo cual es necesario para realizar la evaluación de los modelos de variograma según lo descrito en el capítulo 5. Los parámetros para los diferentes constructores de las clases implementadas son:

- **Coordenadas de los localizaciones:** es una matriz en la que se almacenan las coordenadas de las localizaciones donde se dispone de valores de la variable analizada. Dichas coordenadas deben ser consistentes con un sistema de georeferenciación estándar o un sistema de coordenadas locales definido por el analista. Las coordenadas de las localizaciones pueden extraerse de forma más eficiente empleando un objeto de la clase `ListaObservaciones` y utilizando sus métodos de clase.
- **Valores de las observaciones:** es un vector que almacena los valores conocidos de las realizaciones de la variables analizada, correspondientes a las localizaciones previamente definidas. Pueden extraerse de un objeto previamente construido perteneciente a la clase `ListaObservaciones` empleando los métodos asociados. Igualmente, este vector puede ser ingresado directamente como parámetro de clase y debe estar ordenado de forma que cada valor corresponda a las coordenadas de cada columna en la matriz de localizaciones.
- **Modelo de variograma:** es un objeto de la clase `SuperposicionAnisotropia` que hace posible la evaluación del valor de los variogramas.
- **Georeferenciación de las coordenadas:** es un objeto construido mediante la clase `Georeferenciacion` que indica cual es el sistema de coordenadas que se está usando para localizar los entidades geográficas puntuales con información.
- **Georeferenciación destino:** es un objeto de la clase `Georeferenciacion` que se usa en el caso de que se estén usando coordenadas geográficas para representar la ubicación de las localizaciones. El sistema de georeferenciación destino debe corresponder a un sistema de coordenadas planas y es a éste al cual se proyectan las coordenadas de las localizaciones para fines de cálculo de distancias euclídeas y evaluación de modelos de variograma.

Los parámetros anteriores son comunes para todos los constructores de las clases que implementan los diferentes tipos de krigeados, por lo que en adelante se los nombrará como *Parámetros comunes de krigeado*. Para la descripción de los constructores de las clases de

estimación restantes, se hará referencia a ellos indicando que el constructor correspondiente usa estos mismos parámetros y su descripción es idéntica.

Una vez se ha construido un objeto de las clases `KrigingSim2D`, `KrigingSim3D`, `KrigingOrd2D` y `KrigingOrd3D`, se puede proceder a efectuar la estimación del atributo o variable de análisis en las localizaciones donde no se dispone de información. El método encargado de la estimación del campo raster se ha nombrado como `estimarMalla` y es común para todos los métodos de estimación implementados en las diferentes clases, aunque sus parámetros varían de acuerdo con el tipo de algoritmo. Para los casos de *KS* y *KO* los parámetros del método son:

- **Georeferenciación original y de destino:** son objetos construidos usando la clase `Georeferenciacion` que indican el sistema de georeferenciación de las coordenadas asociadas a las localizaciones con información y el sistema de georeferenciación que se desea para el campo interpolado resultante de la aplicación del método.
- **Metadatos del raster de salida:** son las características que definen el dominio espacial del campo raster de salida del método, a saber, las coordenadas de la esquina inferior izquierda, el tamaño de las celdas en cada una de las tres direcciones coordenadas y el número de celdas (filas, columnas y capas).
- **Media regionalizada:** en el caso en el que se desea realizar las estimaciones utilizando el algoritmo de *KS* debe definirse el valor regional de la media de la variable analizada para el dominio espacial de estimación.

6.4. Algoritmos de estimación que incorporan una tendencia

Las técnicas de krigeado simple y ordinario utilizan los datos de la variable a interpolar para construir cada uno de los estimadores que de ella se buscan. Muchas veces es frecuente contar con alguna o algunas variables relacionadas con la que se está estudiando, las cuales interesa incluir en el proceso de interpolación. Varias son las técnicas que permiten combinar información, entre ellas se han implementado en ESTELLUS el krigeado con una deriva externa y algunas variantes del cokrigeado que veremos más adelante.

6.4.1. Krigeado con deriva externa

El algoritmo de krigeado con deriva externa (*KDE*) considera una función de tendencia definida en cada localización a partir de una variable secundaria $y(\mathbf{u})$, que debe variar continuamente en el espacio. De esta forma, es necesario disponer de un valor de la variable secundaria en cada localización a estimar. En general, se considera que la variable secundaria $y(\mathbf{u})$ está relacionada de forma lineal con la media de la variable primaria $m(\mathbf{u})$, de forma que:

$$m(\mathbf{u}) = a_0 + a_1 y(\mathbf{u}) \quad (6.12)$$

- **Parámetros comunes de krigado:** se definen de manera análoga a los presentados para los algoritmos univariados.
- **Campo raster de derivas:** Es un objeto de la clase `MetaRaster` o `MetaRaster3D` que contiene la información secundaria exhaustivamente muestreada que se desea usar como deriva.

En la implementación del *KDE* se ha contemplado la posibilidad de que la información secundaria no exista en algunas localizaciones. Esto hace que desde el punto de vista de costo computacional, los constructores de las clases correspondientes al *KDE* sean los más demandantes puesto que deben verificar las siguientes condiciones:

- A cada localización puntual de la variable principal debe corresponderle el valor de deriva localizado en la celda más cercana. Para ello, el algoritmo filtra las realizaciones disponibles de la variable principal, de forma que se encuentren incluidas dentro del dominio espacial dado por el campo raster utilizado como deriva.
- Las derivas correspondientes a las localizaciones con información principal no deben ser valores faltantes. Para ello el algoritmo verifica las celdas del campo que se encuentran colocalizadas con la información principal.

Finalmente, se extraen los valores de las derivas incluidas en el dominio espacial y no faltantes, de forma que se pueda construir el sistema lineal de *KDE* descrito en la ecuación 6.14. En las clases `KrigingExtDr2D` y `KrigingExtDr3D`, se incluye el método `interpolarMalla`, en este caso caracterizado por un único parámetro: un objeto de la clase `Georeferenciacion` que indica el sistema de coordenadas del campo raster estimado. En este caso, el dominio espacial y los metadatos del campo de salida son heredados directamente de los correspondientes para el campo de derivas. Cabe aclarar que el campo resultante hereda la georeferenciación del campo de derivas a pesar de que el parámetro de destino para el método `interpolarMalla` no corresponda con el mismo sistema de georeferenciación de las coordenadas para localizaciones de las observaciones de la variable principal.

6.5. Algoritmos de estimación que incorporan información externa

Las técnicas de krigado presentadas hasta aquí sólo se diferencian en la manera en que incorporan la superficie media en un problema de interpolación espacial. En cuanto a la incorporación de información secundaria, los algoritmos de krigado con tendencia requieren que la información secundaria esté disponible en todas las localizaciones a estimar. En muchas situaciones prácticas la información secundaria no está disponible de forma exhaustiva. Para superar dicho obstáculo, se desarrolló la técnica del cokrigado, la cual permite incorporar una o varias variables secundarias en la estimación de un atributo

principal. El cokrigado no se diferencia del krigeado desde un punto de vista algorítmico, su novedad es que permite que los datos secundarios participen directamente en la estimación de la variable principal. Así, desde el punto de vista de la incorporación de atributos externos, el cokrigado hace mejor uso de la información en el sentido de que se consideran tanto la magnitud de la variable secundaria como sus patrones de variación espacial, para la estimación de la variable principal. La cantidad de correlación es capturada por las covarianzas cruzadas y, a diferencia del *KDE*, la variable externa no necesita estar disponible en todos los puntos en los que se desea estimar la variable primaria. La forma general que toma el estimador por cokrigado para una variable $Z_1(\mathbf{u})$ en una localización \mathbf{u} , disponiendo de n_1 realizaciones de la variable principal, denotadas como $z_1(\mathbf{u}_{\alpha_1})$, con $\alpha_1 = 1, \dots, n_1$, es la siguiente:

$$\begin{aligned} Z_{CK}^{(1)*}(\mathbf{u}) - m_1(\mathbf{u}) &= \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1}^{CK}(\mathbf{u}) [Z_1(\mathbf{u}_{\alpha_1}) - m_1(\mathbf{u}_{\alpha_1})] + \dots \\ &\dots + \sum_{i=2}^{N_v} \sum_{\alpha_i=1}^{n_i} \lambda_{\alpha_i}^{CK}(\mathbf{u}) [Z_i(\mathbf{u}_{\alpha_i}) - m_i(\mathbf{u}_{\alpha_i})] \end{aligned} \quad (6.16)$$

donde $i = 2, \dots, N_v$, son los datos secundarios correspondientes a $(N_v - 1)$ variables secundarias; m_1 es la media de la variable principal; m_i son las medias de las variables secundarias y $\lambda_{\alpha}^{CK}(\mathbf{u})$ los coeficientes de ponderación de los datos que intervienen en cada estimación. La localización y el número de los datos secundarios puede ser diferente de una a otra variable y con respecto a los datos principales. Por simplicidad de notación matemática, en adelante sólo se considerará una variable secundaria a menos que se indique lo contrario. Al igual que para krigeado, los coeficientes en la ecuación 6.16, se obtienen resolviendo el sistema de ecuaciones lineales que surge de minimizar la varianza del error (ecuación 6.4) en cada estimación sujeta a la restricción de estimador insesgado (ecuación 6.5). A continuación se describen las distintas variantes del cokrigado implementadas en ESTELLUS.

6.5.1. Cokrigado simple

Si se dispone de datos de una variable principal $Z_1(\mathbf{u})$ y de una variable secundaria $Z_2(\mathbf{u})$, el estimador por *CKS* de la primera viene dado por la siguiente expresión:

$$Z_{CKS}^{(1)*}(\mathbf{u}) - m_1 = \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1}^{CKS}(\mathbf{u}) [Z_1(\mathbf{u}_{\alpha_1}) - m_1] + \sum_{\alpha_2=1}^{n_2} \lambda_{\alpha_2}^{CKS}(\mathbf{u}) [Z_2(\mathbf{u}_{\alpha_2}) - m_2] \quad (6.17)$$

donde m_1 y m_2 son las medias globales respectivas; n_1 es el número de datos primarios y n_2 es el número de datos secundarios. Los pesos $\lambda_{\alpha}^{CKS}(\mathbf{u})$ se obtienen resolviendo el

siguiente sistema de ecuaciones lineales o sistema de *CKS*:

$$\left\{ \begin{array}{l} \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1}^{CKS}(\mathbf{u})C_1(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_1}) + \sum_{\beta_2=1}^{n_2} \lambda_{\beta_2}^{CKS}(\mathbf{u})C_{12}(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_2}) = \\ \qquad \qquad \qquad = C_1(\mathbf{u}_{\alpha_1} - \mathbf{u}), \text{ con } \alpha_1 = 1, 2, \dots, n_1(\mathbf{u}) \\ \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1}^{CKS}(\mathbf{u})C_{12}(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_2}) + \sum_{\beta_2=1}^{n_2} \lambda_{\beta_2}^{CKS}(\mathbf{u})C_2(\mathbf{u}_{\alpha_2} - \mathbf{u}_{\beta_2}) = \\ \qquad \qquad \qquad = C_{12}(\mathbf{u}_{\alpha_2} - \mathbf{u}), \text{ con } \alpha_2 = 1, 2, \dots, n_2(\mathbf{u}) \end{array} \right. \quad (6.18)$$

donde intervienen tanto la covarianza de la variable primaria $C_1(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_1})$ como la de la secundaria $C_2(\mathbf{u}_{\alpha_2} - \mathbf{u}_{\beta_2})$ y la covarianza cruzada $C_{12}(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_2})$. De forma análoga al *KS*, el algoritmo de *CKS* requiere conocer las medias de las variables usadas para la estimación. La expresión para la varianza de la estimación mediante el algoritmo de *CKS* es la siguiente:

$$\sigma_{CKS}^2(\mathbf{u}) = \sigma_1^2 - \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1}^{CKS}(\mathbf{u})C_1(\mathbf{u}_{\alpha_1} - \mathbf{u}) - \sum_{\alpha_2=1}^{n_2} \lambda_{\alpha_2}^{CKS}(\mathbf{u})C_{12}(\mathbf{u}_{\alpha_2} - \mathbf{u}) \quad (6.19)$$

6.5.2. Cokrigado ordinario

El estimador por *CKO* considerando sólo una variable secundaria viene dado por la siguiente ecuación:

$$Z_{CKO}^{(1)*}(\mathbf{u}) = \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1}^{CKO}(\mathbf{u})Z_1(\mathbf{u}_{\alpha_1}) + \sum_{\alpha_2=1}^{n_2} \lambda_{\alpha_2}^{CKO}(\mathbf{u})Z_2(\mathbf{u}_{\alpha_2}) \quad (6.20)$$

en la cual los coeficientes $\lambda_{\alpha}^{CKO}(\mathbf{u})$, resultan de la solución del siguiente sistema de ecuaciones lineales:

$$\left\{ \begin{array}{l} \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1}^{CKO}(\mathbf{u})C_1(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_1}) + \sum_{\beta_2=1}^{n_2} \lambda_{\beta_2}^{CKO}(\mathbf{u})C_{12}(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_2}) + \\ \qquad \qquad \qquad + \mu_1^{CKO} = C_1(\mathbf{u}_{\alpha_1} - \mathbf{u}), \text{ con } \alpha_1 = 1, 2, \dots, n_1(\mathbf{u}) \\ \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1}^{CKO}(\mathbf{u})C_{12}(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_2}) + \sum_{\beta_2=1}^{n_2} \lambda_{\beta_2}^{CKO}(\mathbf{u})C_2(\mathbf{u}_{\alpha_2} - \mathbf{u}_{\beta_2}) + \\ \qquad \qquad \qquad + \mu_2^{CKO} = C_{12}(\mathbf{u}_{\alpha_2} - \mathbf{u}), \text{ con } \alpha_2 = 1, 2, \dots, n_2(\mathbf{u}) \\ \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1}^{CKO}(\mathbf{u}) = 1 \\ \sum_{\beta_2=1}^{n_2} \lambda_{\beta_2}^{CKO}(\mathbf{u}) = 1 \end{array} \right. \quad (6.21)$$

y `CokrigadoSim3D`, para cokrigado ordinario `CokrigadoOrd2D` y `CokrigadoOrd3D`, mientras que para el cokrigado ordinario estandarizado se denominan `CokrigadoEst2D` y `CokrigadoEst3D`. Las seis clases se encuentran ubicadas en el paquete `JGeoST.estimadores.multivariados.objetos`. Es importante notar que los algoritmos de estimación implementados para cokrigado sólo permiten incluir una variable auxiliar para efectos de estimación. Los constructores para las seis clases son prácticamente idénticos, excepto en el sentido de las dimensiones de la información disponible y el algoritmo usado, obviamente. Los parámetros necesarios para construir los objeto de estimación espacial por cokrigado son:

- **Coordenadas de las localizaciones:** son dos matrices en las que se almacenan las coordenadas de las localizaciones en donde se dispone de valores de la variables principal y auxiliar que se desean incluir en el análisis. Dichas coordenadas deben ser consistentes con un sistema de georeferenciación estándar o un sistema de coordenadas locales definido por el analista. Pueden extraerse de dos objetos de la clase `ListaObservaciones` utilizando los métodos de clase correspondientes.
- **Valores de las observaciones:** son dos vectores que almacenan los valores conocidos de las realizaciones de las variables principal y secundaria. Los valores de las realizaciones pueden extraerse de dos objetos previamente construidos pertenecientes a la clase `ListaObservaciones` empleando los métodos asociados. Igualmente, estos vectores pueden ser ingresados directamente como parámetro de clase y deben estar ordenados de forma que cada valor corresponda a las coordenadas de cada columna en la matriz de localizaciones.
- **Modelos de variograma:** son objetos de la clase `SuperposicionAnisotropia` que hacen posible la evaluación del valor de los variogramas necesarios. En este caso se requieren tres objetos de superposición de modelos anisotrópicos: uno para la variable principal, uno para la variable secundaria y otro para el variograma cruzado entre ambas variables. Como paso previo a la estimación de cokrigado, el analista debe asegurarse de que dicha corregeionalización sea válida de acuerdo con los criterios definidos por Christakos (1984) acerca de la definición positiva de las matrices de covarianza.
- **Sistema de georeferenciación de las coordenadas:** se trata de dos objetos de la clase `Georeferenciacion` que indican los sistema de coordenadas que se están usando para localizar las entidades geográficas puntuales con información acerca de las variables principal y secundaria.
- **Georeferenciación destino:** es un objeto de la clase `Georeferenciacion` que se usa en el caso de que las coordenadas de las localizaciones de la información sean geográficas. Dado que el sistema de georeferenciación de destino es aquel al cual se van a proyectar las coordenadas para el cálculo de las distancias euclídeas y evaluar el valor del variograma, éste debe corresponder a un sistema de coordenadas planas.

El método `interpolarMalla` para las clases que implementan cokrigado recibe los siguientes parámetros:

- **Sistemas de georeferenciación original y de destino:** se trata de dos objetos de la clase `Georeferenciacion` que indican respectivamente:
 1. El sistema de referenciación espacial de las coordenadas asociadas con las localizaciones de los datos de las variables principal y secundaria.
 2. El sistema de georeferenciación que se desea para el campo interpolado resultante de la aplicación del método.
- **Metadatos del raster de salida:** son las características que definen el dominio espacial del campo raster de salida del método:
 1. Las coordenadas de la esquina inferior izquierda, el tamaño de las celdas en cada una de las tres direcciones coordenadas
 2. El número deseado de celdas por filas, columnas y capas, los cuales definen la configuración geométrica del campo estimado.
- **Media regionalizada:** en el caso del *CKOE* se requiere especificar el valor de las medias de las variables principal y auxiliar incluidas en el análisis.

6.6. Algoritmos de estimación que incorporan información externa exhaustivamente conocida

Para solucionar un sistema de cokrigado completo como los presentados en las ecuaciones 6.18, 6.21 y 6.24, se requiere tener conocimiento de las funciones de covarianza de las variables principal y secundaria, y la covarianza cruzada entre ellas, si se incorpora una sola variable secundaria. Lo anterior se traduce en un mayor esfuerzo en términos de modelación espacial y tiempo de ejecución en el ordenador, con respecto a las técnicas univariadas, las cuales sólo requieren la función de covarianza de la variable principal. Muchas veces este mayor esfuerzo de cálculo constituye la principal razón por la cual el cokrigado no es utilizado.

Por otro lado, cuando la información secundaria está exhaustivamente muestreada existe una dificultad adicional a la hora de la implementación de un cokrigado completo, que se manifiesta a través de la inestabilidad de las matrices del sistema. Esto se origina debido a la extrema proximidad y fuerte correlación que existe en general entre los datos secundarios, y las grandes distancias y pobre correlación entre los datos principales.

Las razones anteriores, junto con el hecho de que si hay un dato secundario ubicado en la localización que va a ser estimada, éste tiende a atenuar la influencia de los datos vecinos (efecto pantalla), condujo a presentar la alternativa del cokrigado colocalizado (Xu et al., 1992). Se trata de un cokrigado que como información externa retiene sólo

o en términos de correlograma:

$$\rho_{12}(\mathbf{h}) = \rho_{12}(\mathbf{0})\rho_1(\mathbf{h}), \forall \mathbf{h} \quad (6.29)$$

donde $\rho_1(\mathbf{h})$ es el correlograma de la variable principal, $\rho_{12}(\mathbf{h})$ es el correlograma cruzado entre las variables principal y secundaria y $\rho_{12}(\mathbf{0})$ es el coeficiente de correlación lineal entre éstas. Puede decirse que, bajo un modelo de Markov, la covarianza cruzada $C_{12}(\mathbf{h})$ se calcula escalando la covarianza de la variable principal y el factor de escala es el cociente entre $C_{12}(\mathbf{0})$ y $C_1(\mathbf{0})$. Entonces, el estimador por *CKCM* en su forma estandarizada tiene la forma siguiente:

$$\frac{Z_{CKCM}^{(1)*}(\mathbf{u}_0) - m_1}{\sigma_1} = \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1}^{CKCM}(\mathbf{u}) \frac{Z_1(\mathbf{u}_{\alpha_1}) - m_1}{\sigma_1} + \lambda_2^{CKCM} \frac{Z_2(\mathbf{u}_0) - m_2}{\sigma_2} \quad (6.30)$$

donde los pesos $\lambda_{\alpha_1}^{CKCM}(\mathbf{u})$ y λ_2^{CKCM} , son la solución del sistema de ecuaciones lineales *CKCM*, que es el siguiente:

$$\left\{ \begin{array}{l} \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1}^{CKCM}(\mathbf{u}) \rho_1(\mathbf{u}_{\alpha_1} - \mathbf{u}_{\beta_1}) \lambda_2^{CKCM} \rho_{12}(\mathbf{0}) \rho_1(\mathbf{u}_{\alpha_1} - \mathbf{u}) = \dots \\ \dots = \rho_1(\mathbf{u}_{\alpha_1} - \mathbf{u}), \text{ con } \alpha_1 = 1, 2, \dots, n_1(\mathbf{u}) \\ \sum_{\beta_1=1}^{n_1} \lambda_{\beta_1}^{CKCM}(\mathbf{u}) \rho_{12}(\mathbf{0}) \rho_1(\mathbf{u} - \mathbf{u}_{\beta_1}) + \lambda_2^{CKCM} \rho_2(\mathbf{0}) = \rho_{12}(\mathbf{0}) \end{array} \right. \quad (6.31)$$

en este sistema, si $\rho_{12}(\mathbf{0}) = 0$, entonces $\lambda_2^{CKCM} = 0$ y se ignora al dato secundario. Si $\rho_{12}(\mathbf{0}) = 1$, el sistema se transforma en uno de kriging simple con $(n + 1)$ datos, cuyo resultado es $\lambda_2^{CKCM} = 1$ y $\lambda_{\beta_1} = 0$ para todo β , es decir, el estimador es igual al valor de la variable secundaria. El algoritmo de *CKCM* requiere el mismo esfuerzo de implementación que el *KDE*, y tiene la importante ventaja de permitir que el dato secundario participa directamente en la estimación de la variable principal. Además, y a diferencia del *KDE*, el modelo de correogionalización para *CKCM* se construye con base en un parámetro de calibración, como lo es el coeficiente de correlación, lo que hace que la elección de el modelo pueda verificarse a partir de los datos. Sin embargo, el *CKCM* requiere que la variable secundaria sea conocida en todos los puntos donde se quiera estimar la variable principal.

6.6.3. Los objetos computacionales para la estimación mediante cokriging colocalizado bajo un modelo de Markov

Los objetos computacionales que realizan la estimación mediante *CKCM* se han implementado en las clases `CokrigingColMarkov2D` y `CokrigingColMarkov3D`, pertenecientes al paquete `JGeoST.objetos.multivariados.objetos`. Puede incluirse cualquier número de variables secundarias, representadas mediante campos que comparten el mismo dominio

espacial y las mismas características. La restricción anterior es revisada rigurosamente por los constructores de las clases, ocasionando que estos sean altamente demandantes mientras el objeto se construye. Una vez se dispone del objeto computacional de clase, los recursos de ordenador necesarios para la estimación son similares a los requeridos por el algoritmo de *KDE*. Los parámetros para la construcción de los objeto de las clases son:

- **Coordenadas de los localizaciones:** es la matriz que contiene las localizaciones de las entidades geográficas con información para la variable principal. Las coordenadas de las localizaciones pueden extraerse de un objeto de la clase `ListaObservaciones` utilizando los métodos de clase.
- **Valores de las observaciones:** es el vector que almacena los valores conocidos de las realizaciones de la variable principal. Pueden extraerse de un objeto previamente construido perteneciente a la clase `ListaObservaciones` empleando los métodos asociados. Igualmente, este vector puede ser ingresado directamente como parámetro de clase y debe estar ordenado de forma que cada valor corresponda a las coordenadas de cada columna en la matriz de localizaciones.
- **Modelo de variograma:** es un objeto de la clase `SuperposicionAnisotropia` que evalúa los variogramas o covarianzas de la variable modelada, los cuales son transformados internamente en correlogramas, para que las ecuaciones 6.31 sean aplicables.
- **Información auxiliar:** es un vector de objetos de las clases `MetaRaster` o `MetaRaster3D` que contiene los campos de las variables secundarias a ser usadas en el procedimiento de estimación.
- **Estadísticos de las variables auxiliares:** son tres vectores, dos de ellos contienen las medias y las varianzas de las variables auxiliares, mientras que el último contiene las correlaciones lineales entre cada variable auxiliar y la variable principal.
- **Georeferenciación de las coordenadas:** es un objeto perteneciente a la clase `Georeferenciacion` que indica el sistema de coordenadas que se está usando para localizar las entidades geográficas puntuales con información para la variable principal. La georeferenciación de las localizaciones puede extraerse de un objeto de la clase `ListaObservaciones` empleando los métodos creados para ese fin.
- **Georeferenciación destino:** es un objeto de la clase `Georeferenciacion` que se usa en el caso de que se estén usando coordenadas geográficas para representar la ubicación de las localizaciones. El sistema de georeferenciación destino debe corresponder a coordenadas planas y es a éste al cual se proyectan las coordenadas de las localizaciones para fines de cálculo de distancias euclídeas y evaluación de modelos de variograma.

El método `interpolarMalla` sólo puede usarse en el caso de que los campos de las variables secundarias compartan el mismo dominio espacial como se comentó previamente. Los parámetros necesarios para aplicar el método son:

- **Georeferenciación de destino:** es un objeto de la clase `Georeferenciacion` que indica el sistema de georeferenciación que se desea para el campo interpolado resultante de la aplicación del método.
- **Media de la variable principal:** se requiere especificar el valor de la media regional de la variable principal.

6.7. Algoritmos de búsqueda

Las estrategias de búsqueda integradas en ESTELLUS han sido ampliamente usadas en librerías computacionales como `GSTL` (Ying, 2001; Remy, 2004) y `GSLIB` (Deutsch y Journel, 1998). Se trata de los algoritmos de búsqueda por bloques y la búsqueda por áreas vecinas. Los esquemas conceptuales en dos dimensiones se presentan en la figura 6.1.

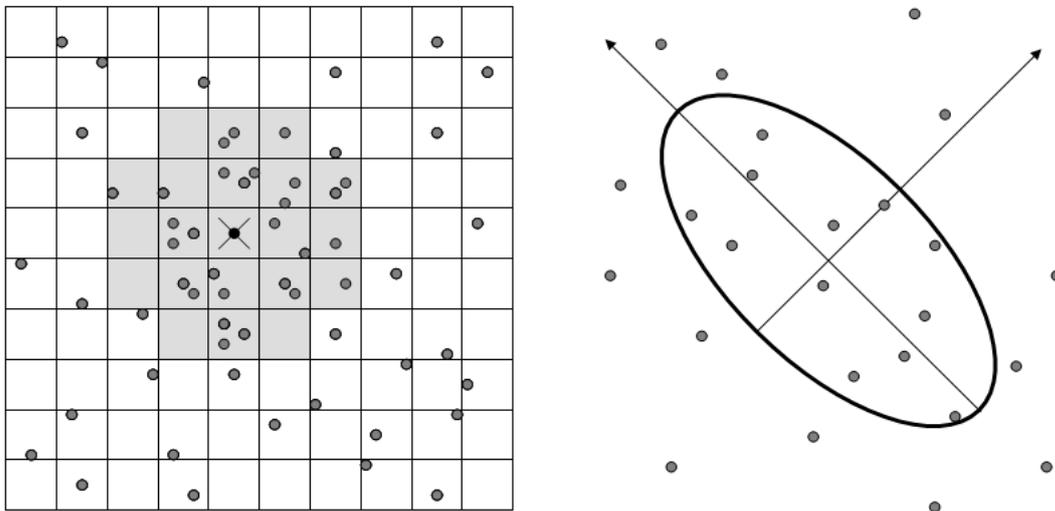


Figura 6.1. Esquema de las estrategias de búsqueda implementadas en ESTELLUS. Búsqueda por bloques (izquierda) y por áreas vecinas (derecha).

A continuación se describen las características generales de las clases implementadas para la ejecución de estas dos estrategias de búsqueda.

6.7.1. Búsqueda por bloques

El algoritmo de búsqueda por bloques se basa en clasificar la información disponible de acuerdo a una red regular de bloques (cuadrados o cubos según se trate de un problema en dos o tres dimensiones, respectivamente). Cuando se realiza una estimación se elige la información que va a intervenir en ella seleccionando los bloques más cercanos al punto siendo estimado hasta completar el número de datos requeridos. En la figura 6.1 (izquierda) el punto negro representa la localización donde se va a realizar la estimación y los bloques de color gris contienen los datos a considerar en ella. Deutsch y Journel (1998) y Journel y Huijbregts (1978) describen en detalle las características operativas del algoritmo.

Para realizar la clasificación por bloques se han creado las clases `OrderingOperations2D` y `OrderingOperations3D`, según se trate de una interpolación en dos o tres dimensiones, respectivamente. Estas clases se encuentran en el paquete de herramientas `JGeoST.estimadores.utilidades`. El constructor de la clase define las características de los bloques de clasificación recibiendo como parámetros el número de fila, columnas y capas de la malla de bloques. Aplicando el método llamado `clasificar` recorre los puntos con información y los asigna al bloque correspondiente más cercano. Los parámetros que recibe este método son:

- **Coordenadas de los puntos con información:** son las coordenadas de la información disponible la cual se desea clasificar por bloques.
- **Observaciones de la variable:** son los valores de las realizaciones de la variable asociadas a las localizaciones con información y pueden extraerse de un objeto previamente construido perteneciente a la clase `ListaObservaciones` empleando los métodos asociados. Igualmente, este vector puede ser ingresado directamente como parámetro de clase y debe estar ordenado de forma que cada valor corresponda a las coordenadas de cada columna en la matriz de localizaciones.

A partir de la clasificación realizada, conociendo las coordenadas de una localización en la que se desea realizar la estimación, pueden extraerse las observaciones en los bloques de acuerdo con una región de búsqueda definida mediante un objeto de las clases `KrigingSearchEstrategy2D` o `KrigingSearchEstrategy3D`. El método de clasificación arroja como resultado una matriz en la cual las tres (o dos, si el análisis es bidimensional) primeras filas son las coordenadas, (*longitud, latitud, elevacion*) o (*este, norte, elevacion*), y la fila final es el valor de la observación del atributo a estimar correspondiente de las observaciones en la región con las cuales se va a realizar la estimación.

6.7.2. Búsqueda por áreas vecinas

Este algoritmo selecciona la información a considerar en cada estimación mediante una búsqueda por proximidad. El área de proximidad está definida por una elipse o un elipsoide para el cual se definen las longitudes de los semiejes y los ángulos que definen su orientación espacial. Para el algoritmo de búsqueda se establece un máximo número de

observaciones con las cuales se realiza la estimación. Así, cuando en el interior del área de búsqueda se localizan más observaciones de las deseadas, el algoritmo realiza un filtro por distancias estructurales. Igualmente, se establece un número mínimo de observaciones, de manera que si no se encuentran dentro de la región de búsqueda, los tamaños de los ejes del elipsoide se dilatan una fracción con respecto a su tamaño anterior. El coeficiente de dilatación se establece por el analista como paso previo al procedimiento de estimación.

Para la construcción de los objetos que realizan la estrategia de búsqueda por áreas vecinas se diseñaron las clases `KrigingSearchStrategy2D` y `KrigingSearchStrategy3D`, para ser utilizadas en estimaciones con dominios bidimensionales y tridimensionales, respectivamente, que se encuentran en el paquete `JGeoST.estimadores.utilidades`. Entonces, un objeto de la estrategia de búsqueda se construye a partir de los parámetro de la elipse o elipsoide y de los puntos con información encontrados dentro del área. Los parámetros que recibe un constructor de éstas clases son:

- **Ángulos de orientación:** son los ángulos azimut, dirección e inmersión que definen la orientación de los ejes principales del elipsoide de búsqueda. Para su definición vale lo visto en la sección 5.8.
- **Parámetros geométricos del elipsoide:** se trata de los tamaños de los semiejes del elipsoide de búsqueda.
- **Límites superior e inferior de puntos dentro de la elipse:** se trata de la cantidad máxima y mínima de observaciones que deben estar contenidas dentro de la región de búsqueda.
- **Factores de dilatación y encogimiento:** el factor de dilatación es una proporción en la cual se aumenta el tamaño de la región de búsqueda en caso de que no se cumpla con la condición del número mínimo de puntos. Por el contrario, el factor de encogimiento es un factor por el cual se reduce el tamaño de la región de búsqueda cuando no se cumple el número máximo de puntos.

En ambas clases existe un único método llamado `filtrar`, el cual se encarga de revisar cuales son las localizaciones disponibles que cumplen con las especificaciones de la región de búsqueda en relación al punto siendo estimado. Los parámetros del método son:

- **Coordenadas de la localización a estimar:** son las coordenadas, de acuerdo con un sistema de georeferenciación estándar para ESTELLUS, de la localización en la cual se desea estimar el valor del atributo. Las coordenadas de las localizaciones pueden extraerse de un objeto de la clase `ListaObservaciones` utilizando los métodos de clase.
- **Coordenadas de los puntos con información:** son las coordenadas de la información disponible la cual se desea clasificar por proximidad.

- **Observaciones de la variable:** son los valores de las realizaciones de la variable asociadas a las localizaciones con información y pueden extraerse de un objeto previamente construido perteneciente a la clase `ListaObservaciones` empleando los métodos asociados. Igualmente, este vector puede ser ingresado directamente como parámetro de clase y debe estar ordenado de forma que cada valor corresponda a las coordenadas de cada columna en la matriz de localizaciones.

El método de clasificación arroja como resultado una matriz en la cual las tres (o dos, si el análisis es bidimensional) primeras filas son las coordenadas, (*longitud, latitud, elevacion*) o (*este, norte, elevacion*), y la fila final es el valor de la observación del atributo a estimar correspondiente de las observaciones en la región con las cuales se va a realizar la estimación.

6.8. Validación cruzada

La validación cruzada es una técnica para evaluar el comportamiento de un modelo de continuidad espacial utilizado para la estimación de un campo aleatorio. Se basa en la estimación del valor de una variable en una localización donde se dispone de medida quitando ésta y utilizando las demás. El proceso se repite para todas las localizaciones de una en una. El análisis se realiza sobre los residuos que surgen de la diferencia entre los valores estimados y los medidos. Los detalles de la técnica de validación cruzada se pueden ver en Isaaks y Srivastava (1989).

6.8.1. Las clases para efectuar la validación cruzada

Se han creado clases para realizar la validación cruzada dentro de cada una de las técnicas de estimación implementadas. Lo anterior se hizo con el objetivo de que el usuario pudiese realizar procedimientos de validación de forma independiente a la ejecución de las estimaciones de los campos. Estas clases reciben como parámetros de construcción los mismos que los constructores correspondientes a los métodos de estimación. Cada clase dispone de dos métodos que efectúan la validación para coordenadas bidimensionales y tridimensionales, respectivamente.

Las clases implementadas para los algoritmos univariados se ubican en el paquete `JGeoST.estimadores.univariados.objetos`, mientras que para los multivariados éstas se localizan en el paquete `JGeoST.estimadores.multivariados.objetos`.

7

Manual del usuario

7.1. Introducción

En este capítulo se presenta un manual de usuario del entorno de aplicaciones gráficas desarrolladas para la ejecución automática de los algoritmos de análisis estructural y de estimación espacial, discutidos en los capítulos 5 y 6. Los despliegues gráficos utilizados para la visualización de campos raster han sido desarrolladas empleando la librería VISAD (Hibbard, 1995, 1998, 2002), mientras que para la visualización y despliegue de los demás gráficos se ha utilizado la librería JFREECHART (JFree, 2005).

7.2. La interfaz principal

Al iniciar la ejecución de ESTELLUS se despliega la interfaz de visualización principal, ejemplos de la cual se muestran en las figuras 7.1, 7.2 y 7.3. En ellas, puede observarse que las opciones de visualización de los campos raster varían de acuerdo con las dimensiones del dominio espacial del problema analizado. La interfaz principal se ha dividido en 7 sectores fundamentales, los cuales han sido delimitados mediante cuadros de color rojo que han sido numerados en las figuras previamente mencionadas. Cada uno de los sectores agrupa un conjunto especializado de herramientas de análisis aplicables a la información disponible o utilidades específicas de visualización de las entidades u objetos geográficos. Las utilidades implementadas en la interfaz para cada uno de los sectores se describen a continuación.

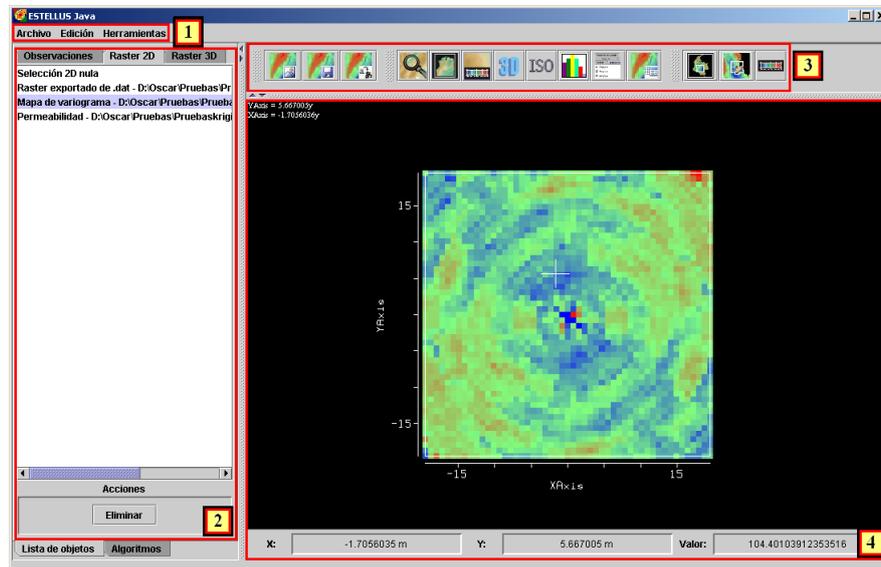


Figura 7.1. Interfaz principal de ESTELLUS. Visualización de un variograma superficial.

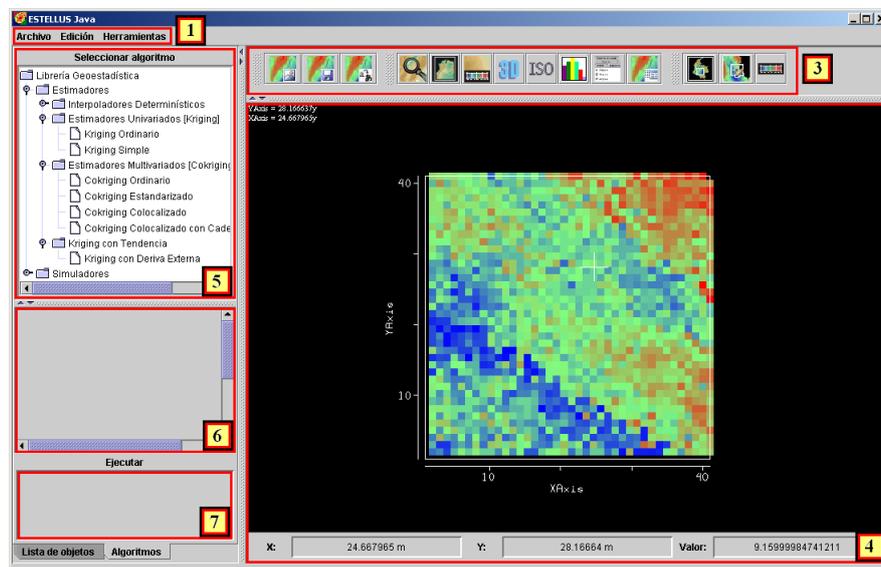


Figura 7.2. Interfaz principal de ESTELLUS. Visualización de un campo raster en dos dimensiones.

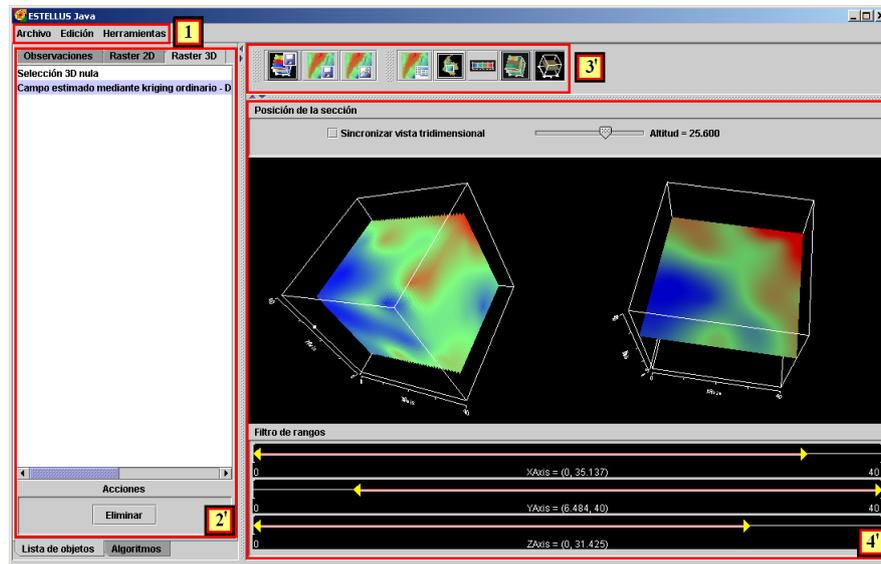


Figura 7.3. Interfaz principal de ESTELLUS. Visualización de un campo raster en tres dimensiones.

7.2.1. Menú de opciones

En las figuras 7.1, 7.2 y 7.3 en la parte superior de la ventana y nombrado mediante el número 1, se muestra el menú general de ESTELLUS, en el cual se han implementado las opciones Archivo, Edición y Herramientas. El opción Archivo ofrece las utilidades de manejo de los archivos de información puntual y los archivo de campos raster. Como norma general, la información puntual se abre mediante archivos de texto plano `*.txt`, con el siguiente formato:

- Primera línea: una cadena de caracteres con el nombre de los datos.
- Segunda línea: sistema de georeferenciación en el cual se presentan las coordenadas de los puntos donde se dispone información (nombre del archivo localizado en el paquete `JGeoST.recurso.georef`).
- A continuación se escriben las coordenadas y la observación de la variables para cada uno de los puntos ordenadas como un vector $[x, y, h, z(\mathbf{u})]$ donde x es la coordenada este o longitud, y es la coordenada norte o latitud, h es la elevación o profundidad desde el nivel de referencia elegido y $z(u)$ es la observación disponible de la variable. Los datos anteriores se escriben en una línea y se separan mediante tabulaciones. En el archivo de texto deben existir tantas líneas como número de observaciones disponibles.

- **Línea final:** En la línea final se escribe la palabra **FIN** para indicar que la lectura de datos ha terminado.

Dicho archivo contiene entonces la información acerca del sistema de georeferenciación de las entidades puntuales, las coordenadas de cada una de ellas junto con la información acerca del valor de la variable representada. ESTELLUS ofrece una interfaz para ingresar de forma interactiva las coordenadas y las observaciones correspondientes a cada uno de los puntos. La información puntual también puede importarse desde archivos de texto en formatos `*.dat`, lo cuales son ampliamente utilizados en librerías geoestadísticas como SGEMS (Remy, 2004) y GSLIB (Deutsch y Journel, 1998). La interfaz encargada de este procedimiento se explica con detalle en la sección 7.4.

La información de los campos raster se lee desde un archivo en formato nativo de ESTELLUS, descrito en la sección 4.4. Los archivos de campos raster se almacenan desde las opciones de almacenamiento de la interfaz de cada algoritmo o se importan desde archivos en formato `ascii` de ESRI o en archivos de texto en formato `*.dat`. En la sección 7.4 se muestran las interfaces construidas para efectuar las operación de importación de campos raster.

Por otro lado, el menú **Edición** ofrece herramientas para cambiar la visualización de los objetos geográficos y para la edición de la información visualizada en los campos raster. Finalmente, el menú **Herramientas** ofrece diversas interfaces gráficas que hacen posible realizar el análisis de la información. Se trata de opciones para el análisis descriptivo de los datos, cálculo de variogramas y ajuste de modelos. En las secciones 7.5, 7.6 y 7.7 se presentan cada una de las interfaces construidas para efectuar las operaciones mencionadas.

7.2.2. Lista de entidades geográficas

En las figuras 7.1 y 7.3, en la parte izquierda se muestra un cuadro rojo que se ha denotado con el número (2), delimitando un panel de múltiples pestañas que contiene las listas de items que permiten manipular la visualización de las entidades geográficas, es decir, los puntos con información y los campos raster bidimensionales y tridimensionales. Cada vez que se abre un objeto geográfico desde el menú **Archivo**, éste queda almacenado en esa lista y puede visualizarse haciendo click sobre la etiqueta correspondiente. Para el caso de las localizaciones con datos se abre una tabla que presenta la información disponible, mientras que en el caso de los campos raster se despliega la visualización gráfica junto con las opciones de las funciones de manipulación implementadas de acuerdo con la dimensión espacial del problema. Para ilustrar al usuario de una forma más clara, en la sección 7.3 se presentan en detalle las diferentes funciones de manipulación para campos raster tridimensionales y bidimensionales.

Con el fin de ahorrar recurso computacional, la memoria **RAM** del ordenador no almacena directamente los objetos geográficos, sino que almacena la ruta correspondiente al archivo donde se almacenan. Entonces, mientras se está desplegando la información es conveniente no transferir los archivos de los objetos que están siendo manipulados actualmente. Si estos

archivos no pueden ser detectados por el programa, se despliega un anuncio de error. Una forma de evitar este inconveniente es crear una carpeta contenedora del trabajo en el disco duro del ordenador, guardar los archivos de los objetos raster en dicha carpeta y no mover los archivos que hayan sido creados mientras ESTELLUS se encuentra en ejecución. Adicionalmente, en la parte inferior del panel múltiple de objetos geográficos se encuentra un botón con la etiqueta **Eliminar** cuya función es borrar de la lista de objetos el que se encuentre seleccionado al momento de pulsarlo. Debe aclararse que la información se borra de la visualización, pero la información de campo raster permanece almacenada en el disco duro del ordenador, por lo cual el usuario debe operar sobre ésta manualmente desde su sistema operativo.

7.2.3. Funciones de manipulación de los campos raster

En las figuras 7.1, 7.2 y 7.3, en la parte superior derecha se observa un conjunto de botones denotados con el número (3). Cada uno de los botones permite el acceso a funciones de manipulación de los campos raster. Son los siguientes:

-  Guardar imagen del campo en formato **jpg**. Esta función está disponible para campos bidimensionales y tridimensionales.
-  Guardar campo raster en disco. Esta función está disponible para campos bidimensionales y tridimensionales.
-  Exportar campo raster a formato **ESRI ascii** raster.
-  Zoom sobre una zona del campo.
-  Restituir vista. Esta función está disponible para campos bidimensionales y tridimensionales.
-  Seleccionar una paleta de colores de una lista predeterminada. Esta función está disponible para campos bidimensionales y tridimensionales.
- **3D** Crear vista tridimensional de un campo bidimensional.
- **ISO** Extraer curvas de isovalores en un campo bidimensional.
-  Estimar histograma y estadísticos del campo raster.
-  Ver una lista con objetos insertados en el despliegue actual. Esta función está disponible para campos bidimensionales y tridimensionales.
-  Ver metadatos del campo raster. Esta función está disponible para campos bidimensionales y tridimensionales.

-  Extraer un campo subconjunto del campo original.
-  Modificar los valores de un campo raster.
-  Personalizar la paleta de colores usada para el despliegue visual. Es una operación de reclasificación de acuerdo con las clases seleccionadas para el histograma.
-  Extraer un subconjunto bidimensional raster desde un raster tridimensional.

Cada una de las interfaces para las funciones implementadas se discuten a fondo en la sección 7.3. Como se observó en lo anteriores items, para el caso de campos raster tridimensionales sólo se implementan algunas de las funciones disponibles para campos bidimensionales, debido a algunas limitaciones de VISAD para la manipulación de los campos tridimensionales.

7.2.4. Panel de visualización de los campos raster

En las figuras 7.1, 7.2 y 7.3 en la parte derecha y denotado con el número (4), se presenta el panel de visualización de campos bidimensionales y tridimensionales. El aspecto y las funcionalidades básicas son diferentes en ambos casos. Por ejemplo, para los campos bidimensionales el valor del atributo representado puede consultarse por *pixel* haciendo click con el botón derecho del ratón y arrastrando el cursor sobre el panel, los valores correspondientes se despliegan en los cuadros de texto ubicados en la parte inferior del panel; mientras que lo anterior no está disponible para campos tridimensionales. Las funcionalidades básicas de despliegue son:

- **Traslación de la visualización del campo:** si se desea arrastrar la visualización del campo, para el caso bidimensional se hace click sobre el panel de visualización y se mueve el puntero del ratón. El cuadro virtual que delimita el dominio espacial del campo raster se mueve de acuerdo con la interacción realizada por el usuario con el ratón. En el caso de despliegue tridimensional debe presionarse la tecla **Ctrl** de forma simultánea con un click de ratón.
- **Rotación de campos tridimensionales:** el cubo que representa el dominio espacial del campo tridimensional se rota haciendo click sobre el panel de visualización y arrastrando de acuerdo a la dirección hacia donde se quiere rotar el cubo.
- **Zoom:** para realizar un zoom se presiona la tecla **shift** de forma simultánea con un click de ratón y se arrastra el cursor hacia adelante o atrás si se desea dar zoom adentro o zoom afuera, respectivamente.

En el caso tridimensional, como se muestra en la figura 7.3, se han implementado dos visualizaciones de los campos. A la izquierda se presenta un cubo que muestra la variación tridimensional del atributo, parámetro o variable analizada, mientras que a la derecha se

presentan un corte bidimensional del cubo a una determinada elevación con respecto a un nivel de referencia determinado previamente por el usuario en los metadatos del raster representado. En la parte superior se dispone de un botón deslizante que ubica la altura a la cual se realiza el corte. También, se dispone de una caja de selección que activa y desactiva la sincronización de la vista del cubo y la vista del corte.

En la parte inferior del panel se presentan tres barras deslizantes que controlan los dominios espaciales de visualización del cubo raster. Deslizandando estas barras pueden obtenerse vistas internas de la variable desplegada. Las barras se encuentran ordenadas desde el eje X hasta el eje Z de arriba hacia abajo.

7.2.5. Árbol de selección de algoritmos de estimación, panel de algoritmo y panel de ejecución

Seleccionando la segunda pestaña del panel de lista de objetos (denominada algoritmos) se accede a las interfaces que permiten el ingreso por parte del usuario de:

1. El tipo de algoritmo de estimación.
2. Las coordenadas y las observaciones en los puntos con información.
3. Los parámetros necesarios para la ejecución de los algoritmos.
4. La selección de las opciones generales en los programas de estimación, como los algoritmos de búsqueda y filtros de vecindad.

El árbol de objetos que se muestra en la figura 7.2, denotado con el número (5), presenta de forma jerárquica los algoritmos implementados en ESTELLUS. En éste se muestran dos categorías principales: **Estimadores** y **Simuladores**, de las cuales sólo se disponen clases Java operativas de la primera. Entrando en la categoría de **Estimadores** se tienen cuatro sub-categorías: estimadores deterministas, estimadores univariados [*kriging*], estimadores bivariados [*cokriging*] y estimadores de kriging con tendencia [*kriging the trend*]. Los algoritmos disponibles por categoría son:

- **Estimadores deterministas:** Se accede a los métodos del inverso cuadrático de la distancia y el método del vecino cercano.
- **Estimadores univariados:** Pueden seleccionarse los algoritmos de kriging simple y el kriging ordinario.
- **Estimadores bivariados:** Se pone a disposición diferentes variaciones de cokriging: cokriging ordinario, cokriging estandarizado, cokriging colocalizado y cokriging colocalizado markoviano.
- **Estimadores de kriging con tendencia:** Se implementa el método kriging con una deriva externa.

Las bases conceptuales y las descripciones de las clases para cada algoritmo se han presentado en el capítulo 6. En la interfaz propiamente dicha, al pulsar click con el botón del ratón sobre el icono del algoritmo deseado, se despliegan las opciones generales correspondientes en el panel denotado con el número (6) (ver figura 7.2) y denominado panel de algoritmos, junto con un texto de encabezamiento que informa al usuario cual de los algoritmos ha sido seleccionado. Evidentemente los paneles que se despliegan al realizar la selección son diferentes para cada algoritmo y cada uno presenta opciones de acuerdo con los requerimientos del procedimiento de estimación correspondiente. Lo detalles de las opciones de los paneles de estimación se describirán con detalle en la sección 7.8. Una vez se han ingresado todos los requerimientos del algoritmo seleccionado puede precederse a ejecutar el algoritmo, para lo cual se pulsa el botón etiquetado como **Ejecutar** que se ubica en el panel denotado con el número (7) (ver figura 7.2) y llamado panel de ejecución.

7.3. Funciones de manipulación, análisis y visualización de campos raster

En la sección 7.2.3 se presentó una descripción breve de las funciones de control y manipulación de los campos raster implementadas en la interfaz principal de ESTELLUS. Dada la importancia que tienen éstas en el análisis e interpretación de los resultados generados en los procedimientos de estimación, a continuación se discuten en profundidad dichas funciones y se presentan las interfaces gráficas de usuario creadas para su utilización.



Figura 7.4. Barra de funciones disponibles para la manipulación de campos raster bidimensionales.

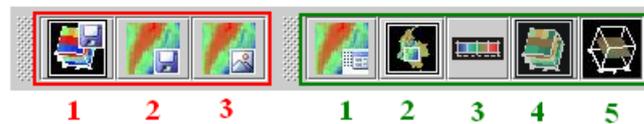


Figura 7.5. Barra de funciones disponibles para la manipulación de campos raster tridimensionales.

En las figuras 7.4 y 7.5 se presentan las barras de herramientas para análisis bidimensional y tridimensional diseñadas para ejecutar dichas funciones de manipulación, las

cuales se encuentran ubicadas en el panel de visualización de campos raster, y se han denotado con el número (3) en las figuras 7.1 y 7.3. El acceso a los cuadros de diálogo para cada función se realiza haciendo click con el ratón sobre los botones correspondientes en la barra de herramientas. Veamos con más detalle cada una de las opciones disponibles.

7.3.1. Guardar imagen de campo

Para guardar una imagen del campo raster que se despliega en el panel de visualización basta con pulsar click sobre el botón , denotado con los números (1) y (3) en las figuras 7.4 y 7.5, respectivamente. Al hacerlo se despliega un selector donde el usuario puede elegir la ruta y el nombre del archivo donde se almacenará la imagen, además puede usar diferentes tipos de formatos para imágenes, por ejemplo *.jpg, *.bmp y *.gif, de acuerdo con sus necesidades. Al finalizar la operación de almacenamiento se obtienen imágenes como las presentadas en la figura 7.6, en la cual, además de la imagen del campo raster, se muestra la paleta de colores junto con la escala de valores utilizada para el despliegue gráfico.

7.3.2. Guardar campo raster seleccionado

Presionando el botón , denotado con el número (2) en la figura 7.4 o, similarmente, el botón , denotado con el número (1) en la figura 7.5, pertenecientes a las barras de herramientas para manipulación de campos bidimensionales y tridimensionales respectivamente, se accede a la función de almacenamiento del campo raster que se encuentra actualmente seleccionado en la lista de objetos de la interfaz principal. Para ello, se despliega un selector de archivo donde el usuario escoge el directorio y el nombre con el cual se desea almacenar el campo. Para finalizar la operación se pulsa click sobre el botón Guardar.

7.3.3. Exportar archivo raster bidimensional actual a formato ascii raster de ESRI

Para acceder a las funciones de exportación de campos bidimensionales desde el formato nativo de ESTELLUS (archivos con extensión *.metaGEO) se hace click en el botón  de la barra de herramientas de visualización denotado con el número (3) en la figura 7.4, para la barra de herramientas de campos raster bidimensionales. Los pasos a seguir para la exportación se explican en detalle en las opciones de archivo, en la sección 7.4.6.

7.3.4. Funciones generales de visualización y manipulación para campos raster

Las funciones generales de visualización de los campos raster se ejecutan desde la barra de herramientas de manipulación, que se presenta en las figuras 7.4 y están numeradas

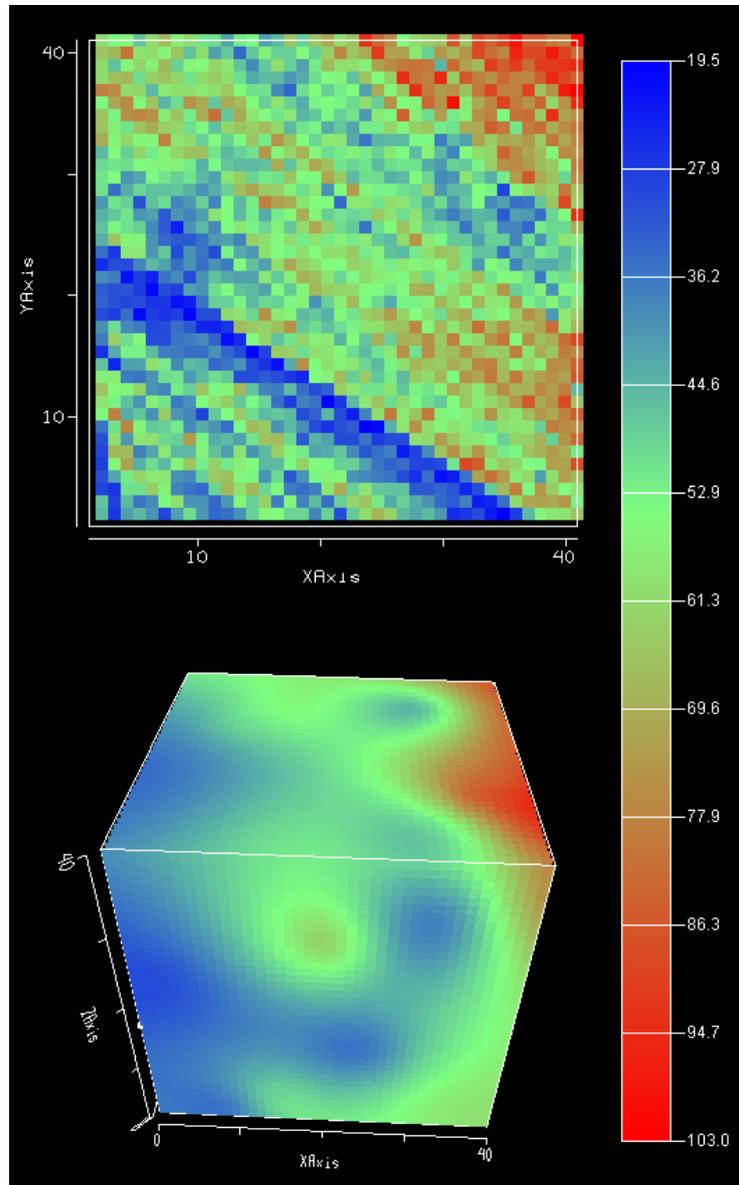


Figura 7.6. Imágenes almacenadas de campos raster bidimensionales y tridimensionales utilizando la herramienta guardar imagen.

desde el (1) hasta el (8) en color verde para campos bidimensionales, y en la figura 7.5 y están numeradas desde el (1) hasta el (5) en color verde para campos tridimensionales. Cada botón tiene una utilidad, las cuales se explican en detalle en los siguientes apartados.

Zoom sobre el campo

El botón , denotado con el número (1) en la figura 7.4, permite realizar un zoom sobre el campo activado actualmente. Para ello se debe presionar el botón y desplazar el cursor sobre el panel de despliegue y se debe definir una región rectangular a la cual se le desea realizar el zoom. Dicho área se indica haciendo click con el botón derecho del ratón en la esquina superior izquierda del recuadro que interese, se mantiene presionado mientras se desplaza el puntero del ratón hasta la esquina inferior derecha del recuadro. Una vez se suelte los botones del ratón, el recuadro seleccionado ocupará toda el área del panel de despliegue.

Restituir campo

Cuando el usuario ha manipulado el campo raster sobre el panel de visualización y desea volver a su visualización original se pulsa el botón , denotado mediante el número (2) en la figura 7.4.

Modificación automática de la paleta de colores

Si el usuario desea modificar las convenciones de colores que está usando para representar los valores de un atributo en el campo raster desplegado actualmente, ESTELLUS pone a su disposición diversas herramientas. En este apartado se describe como seleccionar una paleta de colores de una lista de opciones diseñadas previamente. Para acceder al selector-editor de las paletas de colores disponibles se pulsa el botón , denotado con el número (3) y el color verde en las figuras 7.4 y 7.5.

Como opción principal y por defecto, ESTELLUS tiene una serie de paletas de colores predefinidas que ofrecen diferentes tipos de contrastes, combinaciones continuas y discretas de colores. Cada una de las paletas es caracterizada por su nombre y por la asignación de las proporciones de rojo, verde y azul para cada uno de sus colores. Los componentes de la interfaz para la manipulación de la paleta de colores se muestran en la figura 7.7 y son los siguientes:

- Un campo con el nombre de la paleta actual del mapa, denotado como (i).
- La lista de las paletas predefinidas, denotado como (ii).
- Cuatro curvas que indican el contenido de rojo, verde y azul (RGB) de cada uno de los colores de la paleta y el grado de transparencia de los colores de la paleta, que no es visible pues su valor por defecto es 1, lo que coincide con el límite superior de ésta en la interfaz, contenidos en el panel denotado como (iii).
- Una barra en la cual se muestran cada uno de los 256 colores de la paleta activa, denotada por (iv).

- Un campo de selección que permite ajustar la paleta de colores seleccionada a los datos del mapa activo, denotado como (v).

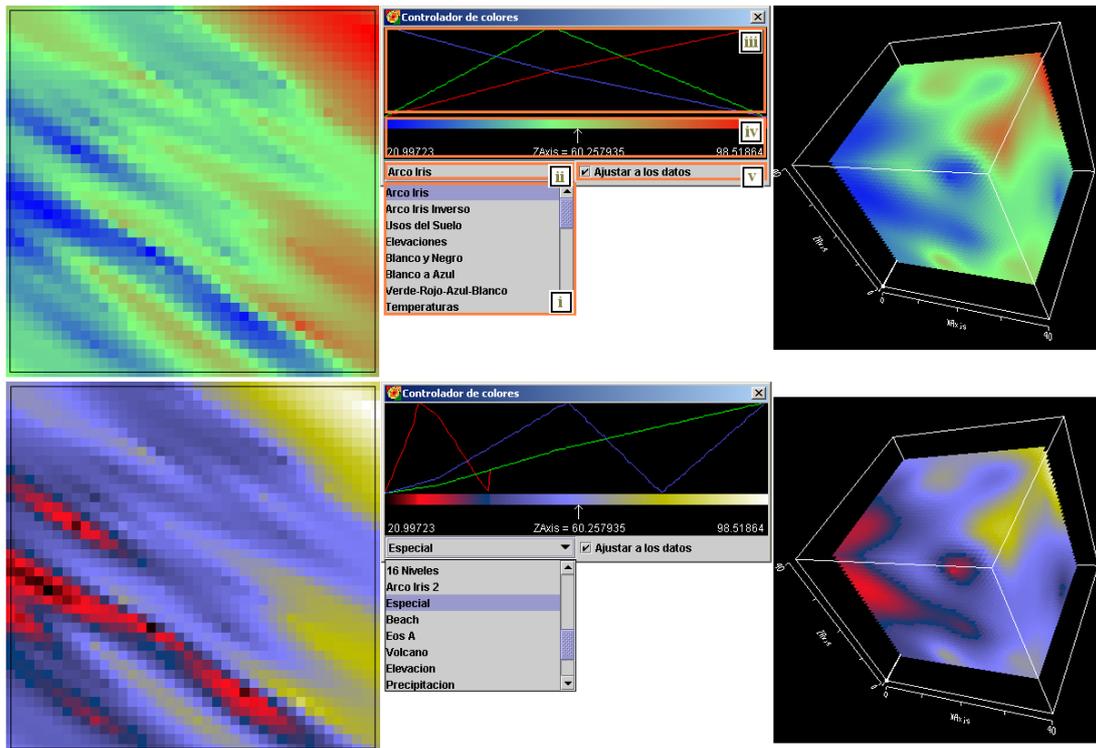


Figura 7.7. Opciones de selección de paletas de colores predefinidas

Existen dos maneras de modificar la paleta de colores activa. La primera consiste simplemente en seleccionar una de las paletas de la lista de predefinidas. La segunda puede ser usada para realizar cambios más detallados, y consiste en modificar la asignación de colores rojo, verde y azul (RGB) de la paleta actual asignando de forma manual los valores de una o varias de las tres curvas de manejo de los colores RGB.

Para realizar lo anterior, el usuario debe ubicarse en el área de las curvas (i) y hacer click con el botón del ratón encima del color de la paleta que quiera modificar, y a una altura igual a la cantidad de rojo, verde o azul que se desea que tenga el nuevo color. Para modificar el siguiente color RGB se hace click con el botón derecho del ratón sobre el panel de comandos (iii) y, automáticamente, se habilita su modificación. El orden de pulsado es: un click para modificar rojo, dos clicks para modificar verde y tres clicks para modificar azul.

Si se quiere modificar la transparencia de la paleta de colores, se hace click con el botón derecho del ratón por cuarta vez sobre cualquier punto del area de las curvas. Se

puede modificar este valor para algún color de la paleta haciendo click a una altura igual al valor de transparencia que se desea. También, si se pretende modificar el nivel de la transparencia para todos los colores de la paleta en el rango de los datos del campo actual, simplemente se mueve el puntero del mouse siguiendo la trayectoria que desea para la curva de transparencia. Todos los cambios realizados sobre la paleta se realizan simultáneamente sobre el despliegue gráfico.

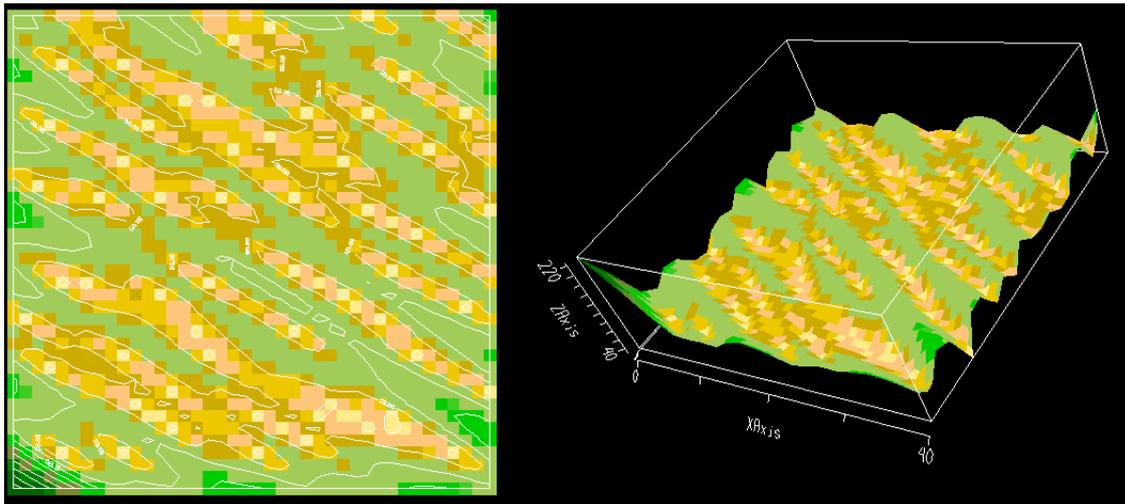


Figura 7.8. Visualización tridimensional y de isolíneas de campos bidimensionales.

Visualización tridimensional de campos bidimensionales

ESTELLUS permite visualizar campos raster bidimensionales de forma tridimensional definiendo una superficie tridimensional. Para acceder a esta herramienta se hace click con el ratón sobre el botón **3D**, denotado con el número (4) y el color verde en la barra de funciones, figura 7.4. Al hacerlo se despliega una visualización tridimensional, como se muestra en la figura 7.8, la cual tiene las mismas funcionalidades proporcionadas por VisAD, comentada en la sección 7.2.4.

Visualización de curvas de nivel sobre campos bidimensionales

ESTELLUS, implementando una funcionalidad integrada en la librería VisAD, permite estimar y extraer automáticamente las isolíneas para la variable representada en un campo raster. Para lograr lo anterior se hace click en el botón **ISO**, denotado por el número (5) en la barra de herramientas de manipulación, mostrada en la figura 7.4. En la figura 7.8 se muestra un resultado típico de efectuar esta operación.

Visualización de los metadatos del campo

Para observar los metadatos del campo raster, es decir, su nombre, sistema de georeferenciación, número de filas, columnas y capas, tamaño de celda, formato de número, etc; debe pulsarse el botón , denotado con los números (8) y (1), de color verde en las figuras 7.4 y 7.5.

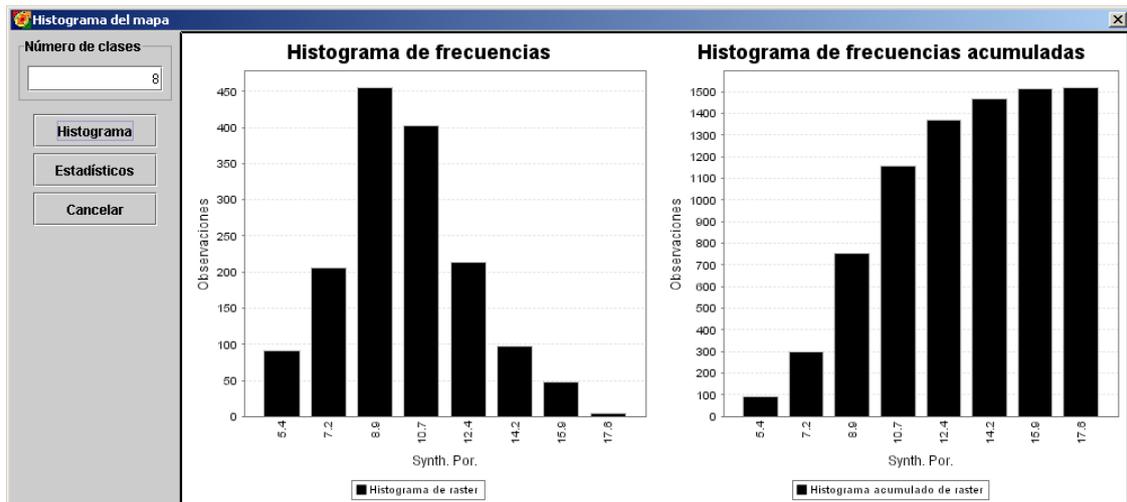


Figura 7.9. Interfaz para la estimación del histograma de un campo raster bidimensional.

Estimación del histograma de frecuencias para un campo bidimensional

En caso en que el usuario quiera personalizar la paleta de colores utilizada para la representación de las imágenes visualizadas, debe estimar el histograma de frecuencias de forma previa puesto que el constructor de la paleta requiere derivar información del histograma. Para ingresar a la interfaz de usuario correspondiente, presentada en la figura 7.9, se pulsa el botón , denotado con el número (6) en la barra de herramientas (ver figura 7.4). En ella, ingresando solamente el número de clases se calcula el histograma e histograma acumulado, junto con algunos estadísticos de la variable representada.

Extraer sub-campo

Esta herramienta sirve para crear nuevos campos mediante la extracción de fragmentos de campos previamente existentes. Para utilizarla, deber pulsarse click sobre el botón , denotado con los números (1) y (2) de colores verde y azul en las figuras 7.4 y 7.5, respectivamente para despliegue gráfico de campos bidimensionales y tridimensionales. Entonces, se seleccionan las coordenadas de la primera esquina ya sea entrándolas manualmente o

haciendo click en el campo raster activo. Lo anterior se repite para la segunda esquina, con lo cual queda definido un nuevo dominio espacial, en el cual quedará definido el nuevo raster. Por último se presiona el botón **Extraer Submapa** para obtener el nuevo campo a partir de la región elegida.

Modificar campo bidimensional

La posibilidad de modificar un campo raster de forma interactiva sólo se encuentra implementada para campos raster bidimensionales. El funcionamiento de la herramienta de edición es muy sencillo. A ella se accede presionando el botón , denotado con el número (2), de color azul en la figura 7.4 y se define el nuevo valor para el pixel que se seleccione utilizando el puntero del mouse sobre el campo raster activo.

Personalizar paleta de colores

Para personalizar la paleta de colores, una vez creado el histograma del campo se presiona el botón , denotado por el número (3) de color azul en la figura 7.4, para que aparezca el cuadro de diálogo mostrado en la figura 7.10, en la que se identifican los siguientes elementos:

- Un cuadro de texto para definir el nombre de la nueva paleta que se va a crear (i).
- Un panel que incluye los límites de cada clase del histograma, que deben ser editados por el usuario (ii).
- Para cada clase se tiene una etiqueta con el color elegido para los valores del mapa dentro de los límites del intervalo correspondiente. Al presionar click sobre la etiqueta aparece un diálogo para elegir el color deseado (iii).
- Un campo de verificación para determinar si se quiere guardar la paleta de colores creada, de tal manera que la próxima vez que se abra ESTELLUS dicha paleta sea una más de las predefinidas (iv).
- Los botones **Aceptar** y **Cancelar**, para seleccionar la combinación de colores o finalizar el procedimiento sin realizar los cambios, respectivamente (v).

Luego de crear la nueva paleta de colores se aplica sobre el mapa activo como se muestra en la figura 7.10.

7.4. Opciones de Archivo

En esta sección se describen detalladamente las funciones disponibles para el manejo de los archivos de entrada y salida de información en ESTELLUS.

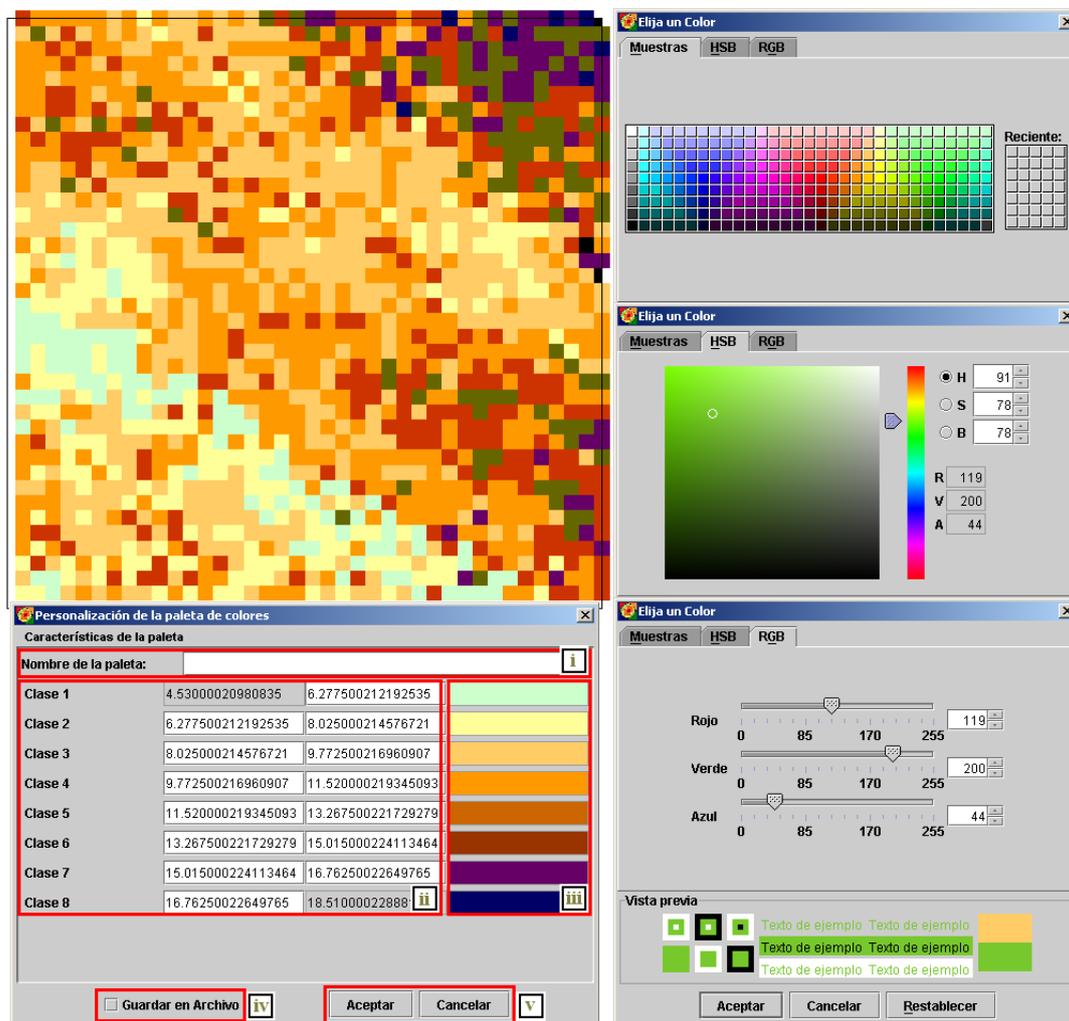


Figura 7.10. Interfaz para la personalización de la paleta de colores de un campo raster.

7.4.1. Abrir un archivo

Para abrir archivos raster se ingresa en el menú principal de la interfaz principal de ESTELLUS, indicado con un (1) en la figura 7.1, en las opciones Archivo y se seleccionan los items Abrir xyh-Z o Abrir raster. Se utiliza el selector de archivo por defecto de Java, como se muestra en la figura 7.11, en la cual se han numerado los componentes que permiten la manipulación de los archivos. El cuadro de opciones (i) permite al usuario elegir la unidad de disco donde se desea realizar la exploración. Los botones que se encuentran a su lado (ii) ejecutan diversas funciones de movimiento en el árbol de directorios, por

ejemplo subir de niveles, ir al escritorio, crear nuevas carpetas o modificar la visualización. En el panel superior (iii) se selecciona el archivo a abrir desde el árbol de directorios del sistema, mientras que en el cuadro rojo inferior (iv) se demarca una caja de selección donde se puede determinar el tipo de archivos que van a ser visibles para el directorio seleccionado de acuerdo con su extensión. Para campos raster archivos tipos *.metaGeo o *.meta3Geo de formato nativo ESTELLUS, y para archivos de información puntual archivos de texto plano *.txt. En la parte inferior del cuadro de diálogo (v) se encuentran ubicados los botones **Aceptar** y **Cancelar**, los cuales ejecutan o cancelan la apertura del archivo, respectivamente. A partir de la información contenida en los archivos se construyen los objetos computacionales que representan a las entidades geográficas. Una vez se leen los archivos, los objetos construidos se almacenan en la lista de objetos y es posible acceder a ellos de la manera descrita en la sección 7.2.2.

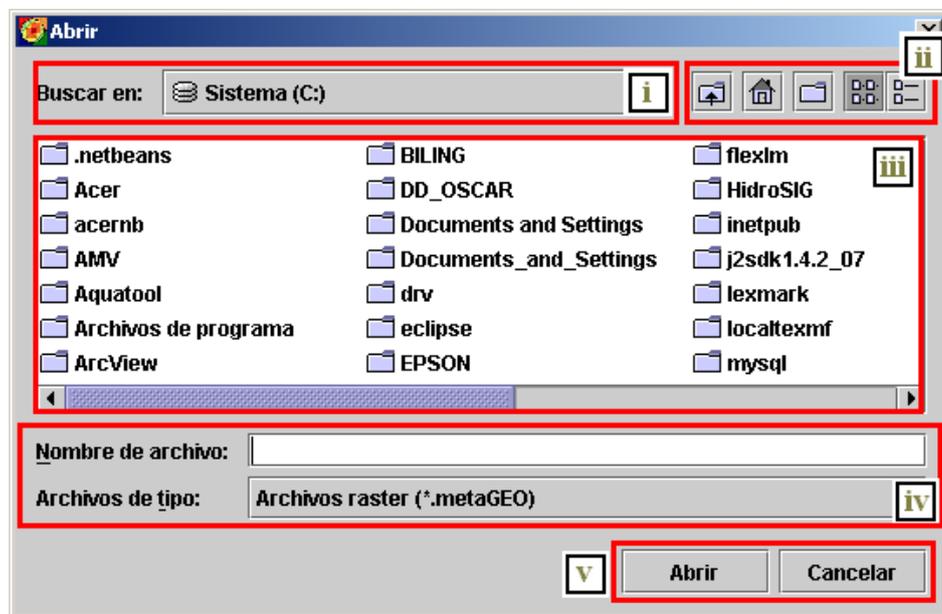


Figura 7.11. Cuadro de diálogo para seleccionar el archivo a abrir.

7.4.2. Entrada manual de información puntual

Otra forma de ingresar la información para crear los objetos que representan las entidades geográficas puntuales es usando la interfaz de ingreso manual de datos que se muestra en la figura 7.12. Para abrirla debe ingresarse al menú principal, (1) en la figura 7.1, en las opciones de **Archivo**, en el ítem **Abrir xyh-Z desde editor**. La interfaz diseñada está compuesta por un panel de definición de información (i), una tabla en la cual se editan los valores de las coordenadas y las observaciones a utilizar en los procedimientos

de estimación (ii) y un conjunto de botones que proporcionan al usuario diferentes tipos de utilidades (iii).

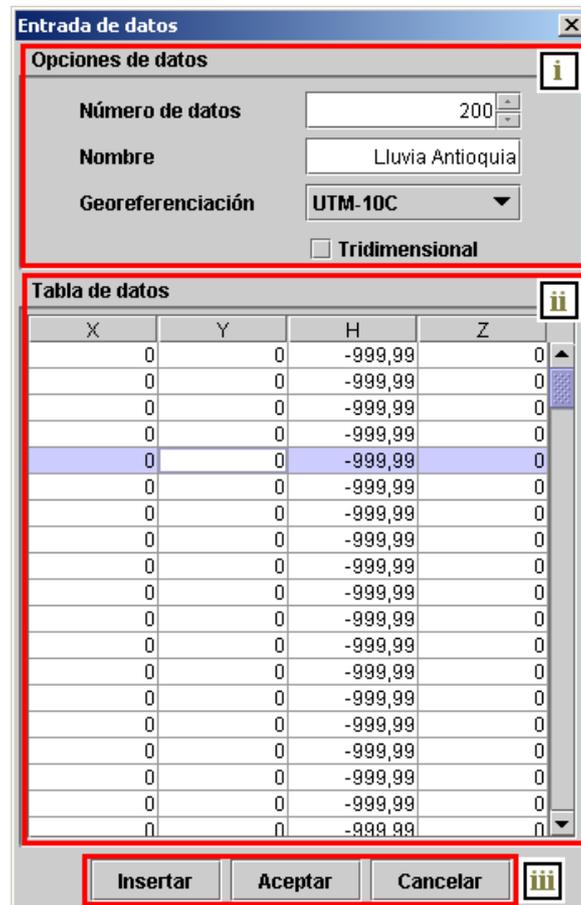


Figura 7.12. Cuadro de diálogo para ingresar la información puntual y de campos raster.

En la parte superior del cuadro de diálogo mostrado en la figura 7.12, se ubica el panel de ingreso de la información, el cual se ha delimitado mediante un cuadro rojo numerado como (i). En éste se ingresa el número de observaciones disponibles, el nombre de la variable a analizar y el sistema de georeferenciación en el cual se están representando las coordenadas de las localizaciones. La tabla indicada con el número (ii) es el editor de la información a ingresar y dispone de varias columnas para el ingreso de las coordenadas y el valor de las observaciones. En la columna X se ingresa las coordenadas este o longitudinales, en la columna Y se ingresan las coordenadas norte o latitudinales, en la columna H se ingresan las coordenadas de la elevación con respecto a un nivel de referencia definido por el usuario y en la columna $Z(\mathbf{u})$ se ingresan los valores de las realizaciones del atributo,

correspondientes a las coordenadas ingresadas en las demás columnas. Si se trata de datos bidimensionales en la columna H aparece por defecto el número -999.99 , que indica que se trata de datos faltantes. Para realizar la entrada de los datos basta con ubicar el puntero del ratón en la celda deseada e ingresar el valor mediante el teclado. El movimiento del cursor dentro de la tabla para ubicar la celda correspondiente se realiza mediante el uso del click en el ratón o las teclas del teclado numérico.

Igualmente, si se desea ingresar una entidad puntual tridimensional, adicional a las definidas previamente, se debe activar el selector con la etiqueta **Tridimensional** en el panel de ingreso de información, con lo cual se elimina la convención de faltantes en las celdas que reciben la información de la elevación. En la parte inferior del cuadro de diálogo se ubican tres botones (iii) etiquetados como **Insertar**, **Aceptar** y **Cancelar**. El botón **Insertar** crea una fila adicional a las existentes en la tabla, susceptible de ser editada. Cuando se ha terminado el ingreso de la información y se desea adicionar el objeto de entidades puntuales dentro de la lista de la interfaz principal se pulsa el botón **Aceptar**, al hacerlo aparecerá un nuevo objeto con el nombre de la variable en la lista de objetos puntuales de la interfaz principal. Si se desea terminar el ingreso de información sin insertar ningún objeto, o simplemente abortar la operación, se pulsa el botón **Cancelar**. Al hacerlo desaparece automáticamente el cuadro de diálogo.

7.4.3. Importar archivos puntuales

Una forma alternativa de obtener archivos de texto en el formato requerido por ESTELLUS, los cuales puedan ser usados efectivamente para crear entidades geográficas puntuales, es importando la información contenida en archivos en el formato GEO-EAS, ampliamente usados en software geoestadísticos como GSLIB (Deutsch y Journel, 1998) y SGEMS (Remy, 2001, 2004). Para acceder a la interfaz que hace posible la operación de importación se debe ingresar en la opción **Archivo** del menú principal, (1) en la figura 7.1, en el ítem **Importar DAT GSLIB...** Cuando éste ha sido seleccionado se despliega la interfaz presentada en la figura 7.13, que se compone de dos paneles (i y ii), una tabla de datos (iii) y un conjunto de botones (iv).

El primer panel (i), llamado **Datos a importar**, contiene dos botones donde se selecciona la localización del archivo que se desea leer y el archivo en el cual se desea escribir los resultados de la importación. Al utilizarlos, las rutas elegidas se escriben en los cuadros de texto asociado con cada botón. En el segundo panel (ii), denominado **Opciones generales**, se presentan cajas de selección donde se seleccionan las columnas del archivo fuente que contienen los datos de las coordenadas X , Y y H y los valores de la variable $Z(\mathbf{u})$ que se desea analizar. También debe seleccionarse el sistema de georeferenciación en el que se representan las coordenadas y debe ingresarse un nombre a la variable. Al seleccionar el archivo fuente se despliega en la tabla de datos toda la información contenida, de forma que el usuario pueda seleccionar las columnas correspondientes que desea integrar al procedimiento de importación. Para finalizar las operaciones de importación y guardar el archivo de texto correspondientes se pulsa el botón **Guardar xyhZ**, ubicado en el panel

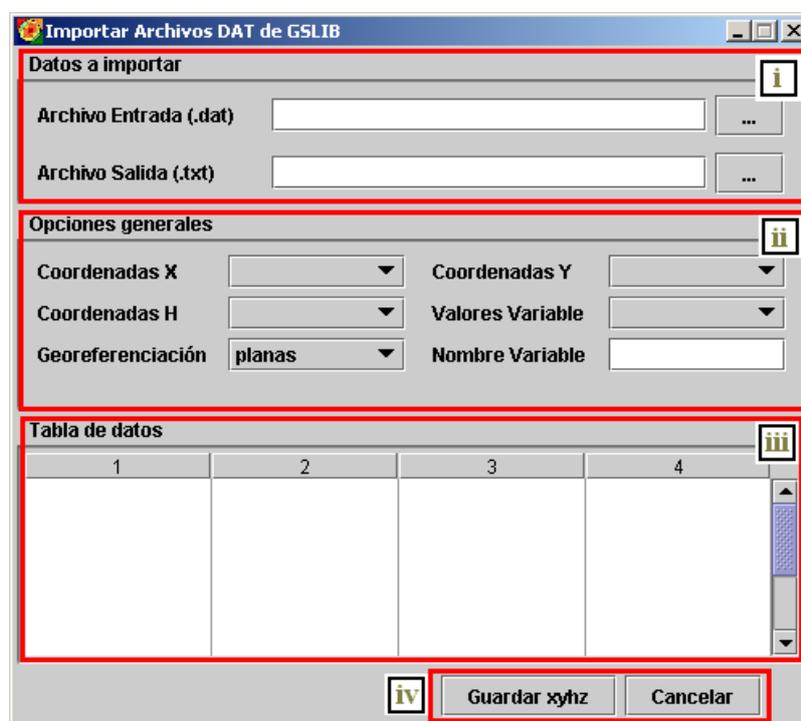


Figura 7.13. Cuadro de diálogo para importar información puntual desde archivos en formato GEO-EAS.

de botones, denotado como (iii). Por el contrario, si se desea cancelar la operación debe hacerse click sobre el botón **Cancelar**, ubicado en el mismo panel. Cuando se termina de importar se obtiene un archivo con el formato descrito previamente en la sección 7.2.1 susceptible de ser leído por ESTELLUS.

7.4.4. Importar campos raster

Los archivos en el formato GEO-EAS también pueden usarse para representar información densamente muestreada, la cual puede ser asociada con archivos de campos raster. Este tipo de representación es bastante usada en librerías geoestadísticas, como GSLIB (Deutsch y Journel, 1998) y SGEMS (Remy, 2001, 2004). En ellos cada línea representa una coordenada en la cual se dispone de información y están ordenadas convenientemente, de forma que se puede construir una malla de información en un dominio espacial determinado. Una vez se han asignado las columnas de la matriz de datos que contienen información de las coordenadas, el algoritmo de importación implementado identifica la celda a la cual corresponde y le asigna el valor de la variable correspondiente. La interfaz implementada en ESTELLUS para el procedimiento de importación se presentan en la figura

7.14. Para acceder a ella debe ingresarse por las opciones de **Archivo** en el menú principal de la interfaz principal de ESTELLUS, (1) en la figura 7.1, en el ítem **Importar Raster DAT...** Al desplegarse, se observa que esta interfaz está compuesta por dos pestañas fundamentalmente, etiquetadas como **Origen-Destino** y **Variable**, cada una de las cuales ofrece diferentes opciones.

La pestaña **Origen-Destino** está formada por dos paneles señalados con cuadros de color rojo. El panel superior (i), llamado **Origen**, contiene la información de los metadatos asociados con el campo raster que va a resultar del procedimiento de importación. Para iniciar dicho procedimiento debe seleccionarse el archivo ***.dat** desde donde se va a extraer la información, el formato o precisión que se desea para el número que representa el valor de la información del campo resultante (**byte**, **short**, **entero**, **flotante** o **doble**), el sistema de unidades en el que se representa la variable, su nombre, la región geográfica que representa el dominio espacial del campo raster y el sistema de georeferenciación al cual se encuentran asociadas las coordenadas en el archivo fuente, denominado como sistema de georeferenciación de origen. En el panel inferior (ii) se selecciona la ruta y el nombre del archivo raster en formato nativo de ESTELLUS, con extensiones ***.metaGEO** o ***.meta3GEO** (es posible importar campos raster tridimensionales), en donde se almacenará la información importada.

Además, debe asignarse el sistema de georeferenciación en el cual se desea representar las coordenadas en el campo raster de salida, el cual no necesariamente debe corresponder al sistema de georeferenciación de las coordenadas en el archivo de origen, salvo en el caso en el que no se esté usando un sistema estándar sino un sistema local personalizado por el usuario. La pestaña **Variable** está compuesta por una tabla de datos (iii), un panel etiquetado como **Dominio y variable del campo raster** (iv) y una caja de selección de la dimensión de los datos a importar (vi). En la tabla de datos (iii) se muestra una vista preliminar de los datos contenidos en el archivo de origen, la cual le permite al usuario identificar las columnas correspondientes a las coordenadas de las localizaciones y la columna que contiene la información de la variable a importar. El panel **Dominio y variable del campo raster** (iv) está dividido en dos pestañas que son:

- **Dominio espacial:** en este panel se encuentran las cajas de texto donde debe ingresarse las características del raster que definen su dominio espacial, es decir, las coordenadas de la esquina inferior izquierda del campo y el tamaño de las celdas que lo conforman, denominadas resolución, como se ha definido previamente en la sección 4.4.
- **Columnas de variables:** permite al usuario seleccionar las columnas en el archivo de origen correspondientes a las coordenadas X , Y y H y $Z(\mathbf{u})$ la variable que se quiere importar al campo raster. Igualmente, hace posible la entrada del número de filas, columnas y niveles del campos raster de salida, los cuales definen los tamaños de las matrices de información a representar.

La información consignada en estas pestañas define el dominio espacial del campo y la variable a representar en él. Por otro lado, si se desean importar campos raster tridimensionales se debe habilitar la caja de selección etiquetada como **Importar campos raster tridimensionales**. Al hacerlo se activan los diferentes cuadros de texto y cajas de selección que permiten especificar las características del campo raster en la tercera dimensión. Finalmente, para completar la operación de importación se pulsa el botón con la etiqueta **Aceptar** (v). Si el procedimiento culminó de forma satisfactoria se despliega un cuadro de advertencia con el texto **Terminó de importar**. Si se desea abortar el procedimiento importación de datos se pulsa el botón **Cancelar** (v) para cerrar el cuadro de diálogo.

7.4.5. Importar desde archivos ascii raster de ESRI

En ESTELLUS se han implementado diferentes utilidades que permiten el ingreso de la información auxiliar de campos raster desde archivos **ascii** raster de ESRI. Esta utilidad ha sido adaptada de la librería de HIDROSIG (Poveda et al., 2007 a), pero se ha modificado para obtener el formato nativo de archivos de campos raster de ESTELLUS, que presenta las características descritas previamente en la sección 4.4. El uso del formato de archivos **ascii** sólo es posible para campos raster bidimensionales.

Para acceder a la aplicación implementada en ESTELLUS se ingresa por el menú principal, (1) en la figura 7.1, dentro de las opciones de **Archivo** en el ítem **Importar Ascii**. Al seleccionarlo, se abre el cuadro de diálogo que se presenta en la figura 7.15, el cual está formado principalmente por dos paneles denominados **ORIGEN** y **DESTINO** etiquetados como (i) y (ii), respectivamente en la figura 7.15, y dos botones, denotados como (iii): **Aceptar** y **Cancelar**.

El panel **ORIGEN** permite la selección del archivo **ascii** que se desea importar al formato nativo de ESTELLUS y el ingreso de los metadatos del campo raster que se va a generar en el procedimiento de importación a saber: (a) el formato o precisión de número con el cual se va a representar la información del raster (**entero**, **flotante**, **byte**, **short** o **doble**), (b) las unidades en las que se representa la variable, (c) el nombre de la variable, (d) la región que está representada en el campo raster y (e) el sistema de georeferenciación correspondiente a las coordenadas espaciales en el dominio espacial del campo raster. Por otro lado, el panel **DESTINO** hace posible la selección del archivo raster bidimensional en formato nativo de ESTELLUS, con extensión ***.metaGEO**, en el cual se almacena la información del campo raster y el sistema de georeferenciación de coordenadas para el campo producto de la importación. Finalmente, para completar la operación se pulsa el botón con la etiqueta **Aceptar** en el panel (iii). Si el procedimiento se culminó de forma satisfactoria se despliega un cuadro de advertencia con el texto **Terminó de importar**; además, si se desea abortar el procedimiento importación de datos se pulsa el botón **Cancelar**, correspondiente al panel (iii), para cerrar el cuadro de diálogo.

7.4.6. Exportar raster bidimensional a formato ascii de ESRI

Se ha creado una herramienta que permite exportar los campos raster estimados a formatos de archivo ascii, compatibles con la mayoría de los sistemas de información geográfica de uso extendido como lo son el ARCGIS, ARCVIEW, IDRISI, etc. Disponer de una herramienta de este tipo en un software que manipula información espacial es sumamente útil, en el sentido de que propicia su uso, no como una herramienta aislada, sino como un entorno que procesa, estima y manipula información geográfica susceptible de ser integrada en un estudio regional, adaptado en un sistema de información geográfica.

La interfaz de usuario implementada que hace posible la exportación de los campos raster estimados a formato `ascii` se presenta en la figura 7.16 y está compuesta por dos paneles: **Origen** y **Destino**, y por dos botones: **Aceptar** y **Cancelar**, los cuales se ubican respectivamente en los paneles denotados como (i), (ii) y (iii) (ver figura 7.16). El panel **Origen** permite seleccionar el campo que se desea exportar, y presenta dos opciones. La primera consiste en exportar un archivo raster previamente almacenado en el disco duro, mientras que la segunda importa el campo raster insertado en la lista de objetos de la interfaz principal que se encuentre seleccionado y que sea visible actualmente. El panel **Destino** permite que el usuario escoja el archivo, de extensión `*.ascii`, en el cual se almacenará la información del campo raster. Finalmente, para completar la operación de importación se pulsa el botón con la etiqueta **Aceptar**. Entonces, si el procedimiento se culminó de forma satisfactoria se despliega un cuadro de advertencia con el texto **Terminó de importar**. Si se desea abortar el procedimiento importación de datos se pulsa el botón **Cancelar** para cerrar el cuadro de diálogo.

7.5. Herramientas para el análisis descriptivo de los datos

En esta sección se describen las interfaces creadas en ESTELLUS para realizar el análisis descriptivo de la información disponible a ser usada en los procedimientos de estimación. Para ello, se implementaron dos interfaces que permiten al usuario realizar el cálculo de los estadísticos univariados y bivariados, las cuales se comentan en detalle a continuación.

7.5.1. Análisis estadístico univariado

En la figura 7.17 se muestra la interfaz implementada en ESTELLUS que permite al usuario realizar un detallado análisis estadístico descriptivo de una muestra de realizaciones de una variable. Se han implementado utilidades para el cálculo de histogramas, histogramas acumulados, funciones de probabilidad acumulada, cálculo de estadísticos, estimación de cuantiles y funciones de probabilidad experimentales. Para acceder a ella se ingresa por el menú **Herramientas** de la interfaz principal de ESTELLUS, (1) en la figura 7.1, en la opción **Análisis Descriptivo, Descripción Univariada**. El cuadro de

diálogo al que se accede está compuesto por un menú de opciones con un único item llamado **Archivo**, demarcado con un cuadro de color azul, y por cuatro paneles delimitados mediante su correspondiente cuadro de color rojo y numerados desde el (1) hasta el (4), que proporcionan las utilidades que se describen en los siguientes apartados.

La opción Archivo

En la opción **Archivo** se selecciona la fuente desde la cual se va a extraer la información de la muestra a analizar. Los algoritmos de cálculo de los estadísticos funcionan con muestras de datos asociadas con un conjunto de localizaciones con información, es decir, cada observación integrada al análisis estadístico debe corresponder a una realización de una variable en una localización. En otras palabras, los algoritmos reciben valores de un atributo asociado con una entidad puntual, que se representa mediante los objetos computacionales descritos en la sección 4.3, listas de observaciones puntuales. Así, las fuentes de información soportadas en esta interface son archivos de texto en el formato nativo de ESTELLUS, como se describió en la sección 7.2.1 de este capítulo, u objetos puntuales insertados en la lista de objetos de la interfaz principal (ver sección 7.4.2).

Para abrir un archivo de texto se escoge la opción **Abrir xyh-Z desde archivo...**, para la cual aparecerá un selector de archivo como el mostrado en la figura 7.11 desde el cual se elige la ubicación del archivo y se crea el objeto computacional de la lista de observaciones (ver sección 4.3.2). Al hacerlo se activan las opciones en la interfaz y pueden llevarse a cabo los análisis deseados. Pueden insertarse tantas variables de análisis como se deseen, puesto que la interfaz guarda las listas de observaciones añadidas, con lo que el usuario, si lo desea, puede volver a calcular los estadísticos para muestras previamente analizadas (ver siguiente apartado). Ahora bien, si se desea estudiar la muestra de las observaciones insertadas en la lista de objeto se selecciona la opción **Abrir desde lista de objetos**, con lo cual se detecta el objeto actualmente activo en la lista de observaciones puntuales en la interfaz principal de ESTELLUS e, igualmente que para la selección de archivos, se activan los campos necesarios para realizar los análisis deseados.

La opción **Archivo** también ofrece la posibilidad de guardar en disco los resultados de los diferentes análisis. Para ello se selecciona la opción **Guardar** y la categoría correspondiente a los resultados que se desean almacenar: histograma, histograma acumulado, función de probabilidad acumulada experimental discreta o probabilidades empíricas; los cuales queda almacenados en archivos de texto a los que pueden accederse usando cualquier procesador de texto convencional.

Panel Opciones de variable

El panel **Opciones de variable**, denotado con el número (1) en la figura 7.17, proporciona las utilidades para elegir la muestra que se desea analizar (puesto que pueden insertarse la cantidad de muestras que el usuario desee analizar, como se comentó en el

apartado anterior). Los análisis pueden realizarse para cualquiera de las columnas que conforman la lista de entidades espaciales, X , Y , H o el valor de la variable.

Así, en el panel (1) se dispone de dos cajas de selección: (a) en la caja de selección llamada **Variable de análisis** se escoge la lista de entidades puntuales, insertadas desde diferentes archivos en formato `*.txt`, que se desea utilizar para realizar los análisis; (b) en la caja de selección llamada **Dato de análisis** se selecciona el campo o columna de la lista de observaciones que se desea analizar. Si se desean eliminar listas de observaciones del conjunto disponible para analizar se pulsa el botón **Eliminar**, de esta forma se borra la lista seleccionada en la caja **Variable de análisis**.

Tabla de datos

En la tabla de datos, denotada con el número (2) en la figura 7.17, se consigna la información de la lista de observaciones asociada con el ítem seleccionado en la caja de selección **Variable de análisis**. Cada fila en la tabla representa las coordenadas de la entidad puntual y el valor del atributo asociado.

Panel Opciones generales

En el panel **Opciones generales**, denotado con el número (3) en la figura 7.17, se ubican las utilidades para el ingreso de los parámetros a los algoritmos computacionales. El panel se encuentra dividido en dos pestañas: **Histograma** y **Cuantiles**. En la primera se ingresan los parámetros requeridos para estimar el histograma: los límites máximos y mínimos de las variables y el número de clases; además, se pueden activar las cajas de selección correspondientes a los análisis a realizar, de acuerdo con los resultados que se desean desplegar al finalizar la ejecución. En la segunda pestaña se define que tipo de cuantiles se desea estimar en el análisis y se pueden especificar diferentes niveles de probabilidad acumulada para los cuales se estimará los valores de los cuantiles asociados.

Panel de botones

En la parte inferior de la interfaz se ubica el panel de botones, denotado con el número (4) en la figura 7.17. El primer botón, llamado **Gráfico QQ Normal...** estima y despliega un diagrama de cuantiles (*QQ-plot*). El segundo botón, **Aceptar**, realiza los cálculos de las opciones señaladas en el panel de **Opciones generales** (3) y despliega los resultados en un cuadro de diálogo individual, como se presenta en la figura 7.18. Si se desea abortar el procedimiento de estimación de estadísticos se pulsa el botón **Cancelar** para cerrar el cuadro de diálogo sin efectuar ningún cálculo.

7.5.2. Análisis estadístico bivariado

En la figura 7.19 se muestra la interfaz implementada en ESTELLUS que permite al usuario realizar el análisis conjunto de dos muestras de observaciones de dos variables. Para

acceder a ella se ingresa por el menú **Herramientas** de la interfaz principal de ESTELLUS, (1) en la figura 7.1, en la opción **Análisis Descriptivo, Descripción Bivariada**. En la librería, se han implementado utilidades para el cálculo de diagramas de dispersión, estadísticos bivariados como la covarianza y el coeficiente de correlación, y el despliegue simultáneos de histogramas para cada variable. El cuadro de diálogo a través del cual se accede a los algoritmos, está compuesto por un menú de opciones con dos items: **Archivo** y **Opciones**, demarcado con un cuadro de color azul, y por cuatro paneles delimitados mediante su correspondiente cuadro de color rojo y numerados desde el número (1) hasta el (4), los cuales proporcionan las utilidades que se describen a continuación.

Los menús Archivo y Opciones

En el menú **Archivo** se seleccionan las fuentes desde donde se va a extraer la información para la variable principal y las diferentes variables auxiliares a integrar en el análisis. De la misma forma que en el caso univariado, los algoritmos implementados para el análisis estadístico descriptivo bivariado funcionan con listas de observaciones, las cuales se construyen desde archivo o se importan desde la lista de objetos insertados de la interfaz principal. Se puede determinar una sola lista de observaciones para la variable principal, mientras que para la variable secundaria se pueden insertar la cantidad de listas de observaciones deseadas por el usuario. Éstas listas definen muestras de variables secundarias según el atributo asociado seleccionado para cada una.

Entonces, para ingresar la variable principal desde archivo se selecciona el item **Abrir xyh-Z principal desde archivo...**, con lo cual se despliega un selector de archivos para seleccionar la ruta deseada. Para ingresar una lista de observaciones de la variable secundaria desde archivo se selecciona el item **Abrir xyh-Z auxiliar desde archivo...**, al seleccionar el archivo que contiene la información se agrega un elemento a la lista de **Variables secundarias** cuya funcionalidad se explica en el siguiente apartado.

Ahora bien, para importar las observaciones cargadas previamente en la lista de objetos de la interfaz principal de ESTELLUS se seleccionan los items **Abrir xyh-Z principal en lista de objetos** y **Abrir xyh-Z auxiliar en lista de objetos**, para las variables principal y auxiliar, respectivamente. La lista de observaciones asignada para la variable principal será aquella que se encuentra seleccionada actualmente de entre los objetos cargados en la interfaz principal, mientras que la lista de observaciones para la variable secundaria se escoge desde una serie de opciones ubicadas en un cuadro de diálogo que se despliega al momento de realizar la inserción de datos. Igualmente, cuando se cargan las listas de observaciones desde la lista de objetos principal, se desactiva la selección de las observaciones insertadas por archivo. Por otro lado, el menú **Opciones** tiene sólo una opción que activa las observaciones de las variables secundarias cuando, previamente, se había elegido cargarlas desde la lista de objetos principal.

Panel de variables

Una vez se ha seleccionado las fuentes de los datos para realizar el análisis, se activan todas las funcionalidades implementadas en la interfaz. También se insertan los nombres de las variables para que el usuario pueda identificar las listas de observaciones con las que se encuentra trabajando. El panel de variables, denotado con el número (1) en la figura 7.19, se compone del cuadro de texto **Variable principal** que muestra el nombre de la variable principal insertada, la lista de **Variables secundarias** que permite la selección de la variable secundaria a integrar en el análisis y el botón **Eliminar** que borra la lista de observaciones seleccionada actualmente en la lista **Variables secundarias**. Los algoritmos de análisis se efectúan utilizando la información de la variable secundaria seleccionada en la lista **Variables secundarias**.

La tabla de datos

En la tabla de datos, denotada con el número (2) en la figura 7.19, se consigna la información de las listas de observaciones asociadas con las variables principal y auxiliares insertadas. Cada fila en la tabla representa las coordenadas de la entidad puntual y el valor de los atributos principal y auxiliar. Cada vez que se selecciona una variable secundaria de la lista **Variables secundarias** se refresca la tabla de datos, mostrando la información asociada.

Panel de parámetros

En el panel de parámetros, denotado con el número (3) en la figura 7.19, se ingresan los parámetros para la construcción de los histogramas de cada muestra: valores máximos y mínimos de las variables principal y secundaria y el número de clases deseado.

Panel de botones

En el panel de botones, denotado con el número (4) en la figura 7.19, se accede al botón **Aceptar**, que efectúa la ejecución de los algoritmos de cálculo y despliega un cuadro de diálogo independiente con los resultados obtenidos, como se muestra en la figura 7.20. De la misma forma, si se desea abortar el procedimiento de estimación de estadísticos se pulsa el botón **Cancelar** para cerrar el cuadro de diálogo sin efectuar ningún cálculo.

7.6. Herramientas para el cálculo de variogramas experimentales

En esta sección se describen las interfaces de ESTELLUS diseñadas para realizar el cálculo de variogramas experimentales y variogramas superficiales.

7.6.1. Cálculo de variogramas directos

En la figura 7.21 se muestra la interfaz creada en ESTELLUS para calcular variogramas experimentales y variogramas superficiales directos. Para ejecutarla se ingresa al menú Herramientas de la interfaz principal de ESTELLUS, (1) en la figura 7.1, en la opción Variogramas experimentales, en el ítem Variograma directo La interfaz Variograma Directo, como se presenta en la figura 7.21, está compuesta de un menú de opciones, que se despliega haciendo click en Archivo, y seis paneles con diferentes herramientas computacionales, los cuales se han señalado mediante cuadros de color rojo y se numeran consecutivamente de (1) a (6). Las utilidades de cada panel se describen detalladamente en los siguientes apartados.

El menú Archivo

En el menú Archivo se encuentran las herramientas que permiten al usuario ingresar la información de las coordenadas y los valores de las observaciones de los atributos a analizar, y guardar los variogramas calculados. De forma similar a las interfaces de análisis estadístico descriptivo, el ingreso de las observaciones se puede realizar desde archivo o desde la lista de objetos puntuales en la interfaz principal de ESTELLUS, de forma que sea posible la construcción del conjunto de entidades puntuales, caracterizadas por sus coordenadas y el valor de un atributo (ver sección 4.3). Para ingresar los valores de las coordenadas y las observaciones del atributo se pulsa sobre el ítem `xyh-Z desde archivo . . .`, con lo cual se despliega un cuadro selector de archivos (ver figura 7.11), en donde se puede escoger la ruta del archivo de texto, en formato ESTELLUS (ver sección 7.2.1), que la contiene. Por el contrario, si se desea utilizar un conjunto de observaciones previamente cargadas en la lista de objetos de la interfaz principal de ESTELLUS se selecciona el ítem `xyh-Z desde lista de objetos` y la interfaz usará para el análisis las observaciones que se encuentren seleccionadas actualmente en la interfaz principal.

También, la opción Guardar variograma del menú Archivo permite al usuario almacenar los variogramas estimados. Si se trata de un variograma omnidireccional se guarda un archivo de texto con la extensión `*.var0`, mientras que para variogramas direccionales la extensión correspondiente es `*.varD`. Ambos tipos de archivos pueden abrirse utilizando cualquier procesador de texto y en su interior se consignan las distancias del centro de clase del variograma, el valor del variograma para cada clase y el número de observaciones por clase. Estos archivos no deben ser modificados pues se usan en otras interfaces para el ajuste de modelos permitidos de variograma y el diseño de modelos anisotrópicos de superposición de variogramas.

Panel de visualización de variogramas

El panel de visualización de los variogramas experimentales estimados, denotado con el número (1) en la figura 7.21, permite la visualización de los variogramas estimados. Si se trata de la estimación del variograma omnidireccional sólo se despliega una ventana

interna con la gráfica del variograma, pero si se trata del cálculo de variogramas direccionales, se despliegan tantas ventanas como direcciones simultáneas se están explorando. Igualmente se despliega una ventana contenedora donde se grafican los variogramas direccionales estimados en un mismo sistema coordenado cartesiano (una especie de diagrama de rosas), en ejes con la misma escala. Esto puede llegar a ser una herramienta bastante útil al momento de identificar las direcciones de máxima y mínima variación, como estimación previa a la ejecución del algoritmo para estimar superficies de variograma.

Barra de herramientas de ejecución

La barra de herramientas de ejecución de algoritmos de variograma experimental, denotada por el número (2) en la figura 7.21, proporciona las herramientas de ejecución de los algoritmos computacionales, los cuales se ejecutan pulsando el botón correspondiente. A continuación se enumeran las funciones de cada botón:

-  Restituir región de búsqueda: presionando este botón se vuelve a la vista original de la región de búsqueda del variograma, mostrada en el panel (3) de la figura 7.21, que se describe en el siguiente apartado.
-  Estimar variograma superficial: al presionar este botón se efectúa el cálculo del variograma superficial de acuerdo con los parámetros establecidos para la región admisible y los parámetros de cálculo ingresados en el panel **Opciones generales**, (4) en la figura 7.21, que se explica en detalle en un apartado posterior. Una vez ejecutado el programa se abre una barra de progreso que muestra el estado del avance de los cálculos. Finalmente, el campo estimado se guarda en disco y se inserta en el panel de visualización de la interfaz principal de ESTELLUS.
-  Calcular variograma experimental: pulsando este botón se estiman los variogramas experimentales deseados de acuerdo con los parámetros de región de búsqueda. Al terminar la ejecución, las gráficas de los variogramas estimados se hacen visibles en las ventanas internas con las gráficas de variograma en el panel (1) en la figura 7.21.
-  Realizar transformada normal: este botón se usa para estimar la transformada Gaussiana de una variable.
-  Estimar transformación indicadora: este botón permite estimar la transformada indicadora de la variable que se analiza, de acuerdo con los parámetros ingresados en la pestaña indicadora del panel **Opciones generales**, (4) en la figura 7.21, y habilita las opciones de estimar el variograma indicador.
-  Ajustar modelo de variograma: una vez se han estimado los variogramas experimentales pueden ajustarse modelos de variograma. Pulsando este botón se accede a

la interfaz que maneja los algoritmos de ajuste semiautomático, que se describieron en la sección 5.7. El funcionamiento de la interfaz de ajuste de variogramas se describe con detalle en la sección 7.7.1. Si se trata de ajustar un variograma direccional, el ajuste se realiza sobre la dirección actualmente seleccionada.

-  Borrar variogramas: pulsando este botón se eliminan las ventanas internas con los variogramas experimentales estimados previamente.
-  Salir: este botón cierra la interfaz y suspende los cálculos computacionales que se lleven a cabo en ese instante.

Panel Región de búsqueda

El panel **Región de búsqueda**, denotado con el número (3) en la figura 7.21, muestra dos esquemas que representan la región del espacio donde deben localizarse las observaciones a ser incluidas en la estimación del variograma direccional, de acuerdo con lo descrito previamente en la sección 5.4, en la figura 5.2. Si se trata de un problema bidimensional el esquema superior muestra las regiones en el plano $X'' - Y''$, mientras que si se trata de un problema tridimensional se activan los esquemas inferiores, en los cuales se representan los cortes en las direcciones de los sistemas coordenados rotados $X'' - Z''$ y $Y'' - Z''$, de acuerdo con las convenciones dadas en la sección 5.8.3. Estas gráficas han sido desplegadas usando la librería VISAD, con lo cual tiene asociadas las funciones descritas en la sección 7.2.4, como lo son el zoom y el desplazamiento espacial del esquema.

Panel Opciones generales

El panel **Opciones generales**, denotado con el número (4) en la figura 7.21, contiene las opciones para establecer el tipo de variograma que se desea estimar y posibilitar la entrada posterior de parámetros para su estimación. Está compuesto por diferentes paneles rotulados, que ofrecen las siguientes opciones:

- **Panel Estimación de variograma.** Ofrece tres cajas de selección para establecer si se trata de variogramas omnidireccionales o direccionales y/o se trata de la variable original o de las transformadas indicadora o Gaussiana. De esta forma, si el cuadro de selección de la opción **Variograma omnidireccional** no se encuentra señalado, la interfaz realiza cálculos de variogramas direccionales puesto que se activan los campos correspondientes en el panel de parámetros (6), de lo contrario se estima el variograma omnidireccional. Igualmente, si el cuadro de selección **Variograma indicador** no está señalado se estima el variograma de la variable original, de lo contrario se estima el variograma indicador. Finalmente, si el cuadro de selección **Variograma T. Gaussiana** se encuentra señalado se estima el variograma de la transformada normal de la variable original, siempre y cuando se haya efectuado dicha transformación empleando el botón previamente comentado en esta sección.

- **Panel Número de direcciones.** Cuando el usuario ha seleccionado un análisis de variograma direccional debe establecer el número de combinaciones de acimutes, direcciones e inmersiones (ver convenciones de dichos ángulos en la sección 5.4) para las cuales se desea realizar las estimaciones. Una vez ingresadas, se crean la misma cantidad de items en la lista de acimutes, direcciones e inmersiones, como se muestra en el panel (5) en la figura 7.21, para posibilitar la asignación de parámetros para cada dirección, como se explicará en detalle en el siguiente apartado.
- **Paneles Origen y Destino.** En el panel **Origen** se escribe el sistema de georeferenciación de los datos ingresados automáticamente al abrir el archivo de datos o seleccionar la lista de observaciones activa; mientras que en el panel **Destino** se establece un nuevo sistema de georeferenciación que será usado para el cálculo de distancias euclídeas.
- **Panel Indicador.** En este panel se introducen los umbrales para el cálculo de la transformada indicadora de una variable principal (los datos duros). Adicionalmente, se presenta una tabla donde se muestra el valor de la transformación para cada valor del atributo representado, de acuerdo con el umbral determinado previamente.
- **Panel Gaussiano:** Este panel contiene una tabla que muestra al usuario el valor de la variable original y su correspondiente transformación normal. Lo anterior se activa dado que la transformada ha sido estimada usando el botón correspondiente contenido en la barra de herramientas de ejecución.

Listado de acimutes, direcciones e inmersiones

La lista de acimutes, direcciones e inmersiones (**Lista de Azi. Dir. e Inme.**), denotada en la figura 7.21 con el número (5), permite el almacenamiento de diferentes conjuntos de parámetros de variograma, asociados con una dirección en el espacio, definida mediante valores únicos de acimut, dirección e inmersión, de modo que la interfaz calcule los variogramas experimentales correspondientes. Lo anterior hace posible la estimación simultánea de tantos variogramas direccionales como lo desee el usuario.

El panel de ingreso de parámetros

El panel de parámetros, denotado con el número (6) en la figura 7.21, tiene un funcionamiento que está relacionado estrechamente con la lista de acimutes, direcciones e inmersiones, puesto que el ingreso de parámetros se realiza para cada conjunto acimut-dirección-inmersión. El ingreso se realiza seleccionando un componente de la lista de acimutes, direcciones e inmersiones, digitando en los campos de parámetros (acimut, dirección, inmersión, tamaño de lag, tolerancia angular, ancho de banda y distancia máxima) el valor correspondiente. Para ingresar los parámetros en cada componente de la lista se hace click con el ratón sobre cada elemento y se repite el ingreso en los campos correspondientes. Si se desea revisar los valores de parámetros ingresados previamente para

una componente basta sólo con seleccionar el ítem deseado y los valores aparecen automáticamente en el campo asociado. Similarmente, si se desea estimar un variograma omnidireccional se definen los parámetros generales en los campos que permanecen activos. Al terminar de ingresar los conjuntos de parámetros para los acimutes, direcciones e inmersiones establecidas, se puede proceder a ejecutar las funciones implementadas en la barra de herramientas de la interfaz, como se discutió en esta sección.

7.6.2. Cálculo de variogramas cruzados

En la figura 7.22 se presenta la interfaz creada para estimar los variogramas experimentales y variogramas superficiales cruzados usando ESTELLUS. Para ejecutarla se ingresa por el menú **Herramientas** de la interfaz principal, (1) en la figura 7.1, en la opción **Variogramas experimentales**, en el ítem **Variograma cruzado . . .**. La interfaz **Análisis de Variograma Cruzado**, como se presenta en la figura 7.22, está compuesta por un menú de opciones, delimitado mediante un cuadro de color azul, y seis paneles con diferentes herramientas computacionales, los cuales han sido señalados mediante cuadros de color rojo y se numeran consecutivamente del (1) al (6). Las utilidades de cada panel son similares a las presentadas en la sección anterior, con lo cual en los siguientes apartados sólo se discutirán las variaciones a las herramientas existentes y las utilidades no existentes en los análisis univariados.

Los menús Archivo y Opciones

El menú **Archivo** permite al usuario el ingreso de la información de las variables principal y auxiliar que van a ser utilizadas en las estimaciones de los variogramas cruzados. El procedimiento es similar al presentado en la sección 7.6.1, con la diferencia de que existen ítems diferenciados para ambas variables, los cuales habilitan la entrada de los datos asociados, así: **xyh-Z principal desde archivo . . .**, **xyh-Z secundaria desde archivo . . .**, **xyh-Z principal desde lista de objetos** y **xyh-Z secundaria desde lista de objetos**. Desde **archivo** es posible ingresar tantas listas de observaciones para diferentes variables secundarias como el usuario desee. Adicionalmente, estas listas quedan almacenadas en memoria RAM y, para efectos de cálculo, la variable secundaria actual se selecciona en la lista **Seleccionar variable secundaria**, cuyas funciones se describirán detalladamente en apartados siguientes.

Cuando se desee ingresar una lista de observaciones de la variable secundaria desde la interfaz principal de ESTELLUS, la lista de datos asociada debe seleccionarse desde un cuadro de diálogo que se abre al hacer click en el ítem respectivo. Idénticamente que para la interfaz para variogramas directos, la variable principal que se adopta para los cálculos desde la lista de objetos será aquella que se encuentra señalada en el momento. El menú **Opciones** permite al usuario reactivar las listas de observaciones insertadas desde **archivo**, dado que en algún momento se haya optado por ingresar la variable secundaria desde la lista de objetos de la interfaz principal de ESTELLUS.

Para los variogramas cruzados también se dispone la opción de almacenar los resultados en un archivo. Para acceder a ella se selecciona la opción **Guardar variograma** del menú **Archivo**. Si se trata de un variograma omnidireccional se guarda un archivo de texto con la extensión `*.var0C`, mientras que si se trata de variogramas direccionales la extensión correspondiente es `*.varDC`. El formato interno del archivo de texto es similar al generado al guardar los variogramas directos, tal y como se expuso en la sección 7.6.1.

Panel de visualización de variogramas

El panel de visualización de los variogramas experimentales estimados, denotado con el número (1) en la figura 7.22, tiene las mismas utilidades que la expuestas para el cálculo de los variogramas directos en la sección 7.6.1. Su función simplemente es presentar las ventanas internas con los variogramas cruzados estimados.

Barra de herramientas de ejecución

Para el variograma cruzado, la barra de ejecución, denotada por el número (2) en la figura 7.22, además de los botones presentados para el variograma directo en la sección 7.6.1, ofrece al usuario las siguientes utilidades:

- 
 Calcular diagrama de dispersión: una vez se hayan ingresado las listas de observaciones para la variable principal puede calcularse el diagrama de dispersión entre la variable principal y una de las variables secundarias consideradas. El diagrama se dibuja sólo para las observaciones correspondientes a la misma localización espacial. Cuando finaliza la ejecución del algoritmo se despliega el gráfico del diagrama de dispersión, mostrado en la figura 7.22 y denotado mediante el número (7).
- 
 Calcular la transformada indicadora de la variable secundaria: cuando se desea realizar una estimación de campos de probabilidad de no excedencia usando cokrigado indicador se hace necesario efectuar la transformación indicadora de la información secundaria. Los detalles conceptuales de lo anterior se han descrito en la sección 4.6. Pulsando click sobre este botón se efectúa la transformación indicadora secundaria, siempre y cuando se haya realizado previamente la transformación indicadora principal.
- 
 Calcular variograma de residuos: calcula los variogramas de la variable residual para las transformaciones indicadoras, definida de la forma $R(\mathbf{u}, z_k) = I(\mathbf{u}, z_k) - y(\mathbf{u}, z_k)$, siendo $I(\mathbf{u}, z_k)$ la transformación indicadora de la variable principal y $y(\mathbf{u}, z_k)$ la transformación indicadora de la variable auxiliar, en la localización \mathbf{u} con un umbral z_k . La modelación de este variograma es necesaria para la implementación del cokrigado indicador, tópicamente no tratado en este trabajo. Los detalles conceptuales de este algoritmo se presentan en Goovaerts (1997).

Panel Región de búsqueda

El panel **Región de búsqueda**, denotado con el número (3) en la figura 7.22, tienen las mismas funcionalidades comentadas en la sección 7.6.1.

Panel Opciones generales

En el panel opciones generales, denotado con el número (4) en la figura 7.22, se encuentran las opciones que el usuario puede utilizar para determinar el tipo de variograma experimental cruzado que desea estimar y, de esta forma, proceder a la entrada de parámetros. Está compuesto por cuatro pestañas que contienen las siguientes utilidades:

- **Pestaña Variables:** en este panel se escoge el tipo de variograma que se desea estimar: omnidireccional, direccional, indicador, etc. Similarmente a lo presentado en la sección 7.6.1, se selecciona el número de variogramas direccionales deseados (cada uno de ellos tiene asociado un acimut, una dirección y un ángulo de inmersión, junto con los demás parámetros de cálculo de variograma direccional, según se describió en la sección 7.6.1) y la lista de observaciones asociadas con la variable secundaria a ingresar en los análisis. Para activar las opciones que permiten realizar los cálculos debe haberse seleccionado una variable secundaria.
- **Pestaña Georeferenciación:** en este panel se escribe el sistema de georeferenciación para las coordenadas de las localizaciones de las observaciones disponibles de las variables principal y auxiliar. También se determina cual será el sistema de coordenadas plano en el cual se realizarán los cálculos de las distancias euclídeas.
- **Pestaña Indicador principal:** en ella se ingresan los parámetros necesarios para el cálculo de la transformación indicadora de la variable principal, necesario para estimar la transformación indicadora secundaria y obtener una función de probabilidad que pueda condicionar los indicadores principales. Para realizar lo anterior el usuario debe ingresar el valor del umbral principal. Al realizarse la transformación usando el botón correspondiente en la barra de herramientas, los valores de la transformada para cada realización de la variable principal se presentan en la tabla asociada en la pestaña.
- **Pestaña Indicador secundaria:** en ella se ingresan los parámetros necesarios para el cálculo de la transformación indicadora secundaria. Debe ingresarse el número de clases que representarán la función de probabilidad acumulada que condicionarará la transformación indicadora de la variable principal. Una vez se haya realizado el cálculo de la transformación los valores de los indicadores secundarios para las realizaciones disponibles se calculan pulsando el botón correspondiente en la barra de herramientas. Posteriormente, los resultados de las transformaciones se presentan en la tabla asociada al panel.

La Lista de acimutes, direcciones e inmersiones

La Lista de acimutes, direcciones e inmersiones, denotada en la figura 7.22 con el número (5), tiene las mismas funcionalidades presentadas para la interfaz *Análisis de Variograma Directo*, presentada en la sección 7.6.1.

El panel de ingreso de parámetros

El panel de parámetros, denotado con el número (6) en la figura 7.21, ofrece los campos de entradas a los parámetros de estimación del variograma cruzado, su funcionamiento es idéntico al panel con el mismo nombre en la interfaz *Análisis de Variograma Directo*, presentada en la sección 7.6.1.

7.7. Herramientas para el ajuste de modelos de variogramas

ESTELLUS ofrece al usuario dos herramientas para el ajuste de modelos de variogramas:

1. Una interfaz que permite ajustar un modelo a un variograma experimental, de forma manual o automática mediante las técnicas de mínimos cuadrados ordinarios y mínimos cuadrados ponderados, para las cuales se presentó un marco conceptual breve en la sección 5.7.
2. Una interfaz que permite la construcción interactiva de un modelo de variograma anisotrópico o una superposición de modelos anisotrópicos, disponiendo únicamente de variogramas experimentales en las direcciones principales de anisotropía.

En las siguientes secciones se presentan los detalles del funcionamiento de ambas herramientas.

7.7.1. Ajuste de modelos a variogramas experimentales

Dentro del conjunto de herramientas de análisis espacial incluidas en la librería ESTELLUS, se encuentra una interfaz que permite realizar ajustes manuales o automáticos de modelos a un variograma experimental previamente estimado. Un ejemplo de la ventana correspondiente se presenta en la figura 7.23. Para acceder a la utilidad computacional desarrollada se ingresa al menú *Herramientas* de la interfaz principal, (1) en la figura 7.1, en la opción *Ajuste de variogramas...* Una forma alternativa de acceder a la interfaz es utilizando las interfaces de cálculo de los variogramas experimentales directos y cruzados, haciendo click el botón de ajuste de variograma en la barra de herramientas, como se describió en la sección 7.6. La interfaz está compuesta por cinco paneles, delimitados por líneas rojas y denotados con su número correspondiente en la figura 7.23, y un menú de opciones, cuyas funciones se describen en detalle en los apartados siguientes.

El menú Archivo

El menú **Archivo**, que se encuentra en la parte superior de la interfaz, se señala mediante un cuadro de color azul en la figura 7.23 y permite al usuario el ingreso de los archivos de variogramas experimentales estimados, que han sido previamente almacenados en disco como se comentó en la sección 7.6. Una vez se ha finalizado con el procedimiento de ajuste del modelo los resultados pueden almacenarse en disco duro en un archivo de texto mediante la opción **Guardar variograma...** Accediendo a ella se debe escoger la ruta del archivo con extensión ***.Vmod**.

Panel de parámetros del modelo a ajustar

En la interfaz se ha implementado un panel para posibilitar el ingreso de los parámetros del modelo o los modelos, en caso de querer ajustar una estructura imbricada, como se muestra en la figura 7.23, en el cuadro rojo marcado con el número (1). Los campos incluidos en este panel permiten al usuario realizar un ajuste manual interactivo del modelo de estructura imbricada de variograma. Si se desea incluir el efecto pepita en la estructura imbricada se inserta el valor correspondiente en el campo **Efecto pepita**, de lo contrario el valor a ingresar es cero. Antes de definir los valores de los parámetros para las estructuras debe ingresarse el número de estructuras que se desea superponer (sin considerar al efecto pepita) en el campo **Número de estructuras**.

Por ejemplo, si se desea ajustar una estructura imbricada compuesta por un efecto pepita y un modelo esférico con rango a y meseta b , debe ingresarse el número 1. En este momento se puede proceder a ingresar los valores de los parámetros para cada estructura, para lo cual se selecciona la estructura a la que se le desea modificar los parámetros en la lista **Estructuras imbricadas** y se procede a ingresar los valores correspondientes en los campos **Modelo**, **Meseta** y **Rango**. Esto se repite para todos los modelos ingresados en la estructura imbricada. Cada vez que se selecciona un modelo perteneciente a la estructura a ajustar, los campos anteriores se refrescan mostrando el último valor ingresado para éste.

Gráfico de variograma experimental y modelo ajustado

En el panel demarcado mediante un cuadro rojo y denotado con el número (2) en la figura 7.23, se muestra un gráfico del variograma experimental junto con el modelo actual ajustado de acuerdo con los parámetros insertados en el panel (1). La línea negra representa el variograma experimental, mientras que la línea azul representa el modelo ajustado actualmente. Esta gráfica se ha construido utilizando la librería **JFRECHART**, por lo cual se dispone de las utilidades descritas brevemente en la sección 3.3.3.

Panel Informe de resultados

La tabla **Informe de resultados**, denotada con el número (3) en la figura 7.23, muestra al usuario los parámetros actuales para cada uno de los modelos de variograma que

componen la estructura imbricada que se está ajustando al variograma experimental. Adicionalmente, en la parte inferior se ha ubicado un panel denominado **Información precedente**, que contiene un informe de los ajustes previos y los botones de ejecución y salida del algoritmo de ajuste automático. Al hacer click sobre el botón **Aceptar** se realizan los procedimientos automático del ajuste de la función del variograma y, una vez terminada la ejecución, los campos correspondientes a los parámetros de ajuste se refrescan junto con la gráfica de ajuste en el panel (2). Si se desea terminar las operaciones y cerrar el cuadro de diálogo, se pulsa el botón **Cancelar**.

Panel de parámetros para el ajuste automático

Los botones y campos de texto para ingresar los parámetros requeridos por el algoritmo de ajuste automático de modelos de variograma, descritos en la sección 5.7, se localizan en el panel delimitado por una línea roja y denotado por el número (4) en la figura 7.23. El usuario debe seleccionar el tipo de algoritmo con el que desea realizar el ajuste: mínimos cuadrados ordinarios o ponderados, el dominio de distancias para el cual quiere efectuar el ajuste en el campo de texto **Distancia máxima** y los parámetros del algoritmo *Nelder-Mead* implementado para el ajuste automático: (i) valores iniciales de los parámetros, (ii) pasos iniciales de deformación del simplex y (iii) parámetros de parada del algoritmo. Para hacer efectivo el ingreso de los parámetros del algoritmo de ajuste deben pulsarse los botones adyacentes al campo de ingreso correspondientes a cada parámetro. Los valores actuales para los parámetros se consignan en la tabla **Informe de resultados**.

Panel de función objetivo

En el panel de función objetivo, denotado con el número (5) en la figura 7.23, se encuentran dos campos de texto que informan al usuario del valor de la función objetivo alcanzada en el modelo ajustado por mínimos cuadrados y el modelo actual, dado por los parámetros ingresados en el panel (1). Al momento de terminar la ejecución del algoritmo de ajuste automático ambos valores son iguales, pero si el usuario encuentra inconsistencias con los parámetros ajustados los parámetros pueden volverse a ingresar, con lo cual el valor de la función objetivo actual cambia. De esta forma se puede visualizar las desviaciones obtenidas haciendo un ajuste manual con respecto al valor óptimo obtenido en el ajuste automático.

7.7.2. Creación de modelos anisotrópicos

ESTELLUS dispone de una interfaz gráfica que permite al usuario crear superposiciones lineales de modelos de variograma anisotrópicos, como se muestra en la figura 7.24. Para acceder a ella se ingresa al menú **Herramientas** de la interfaz principal, (1) en la figura 7.1, en la opción **Evaluar regionalización...** La ventana de diálogo está formada por un menú, ubicado en la parte superior y señalado mediante un cuadro azul, y cuatro paneles de herramientas delimitados mediante cuadros rojos y denotados con

su número correspondiente. En los siguientes apartados se describen las funcionalidades computacionales implementadas en esta interfaz.

Los menús Archivo y Opciones

El menú principal de la interfaz **Ajuste de regionalización** se compone de dos submenús: **Archivo** y **Opciones**. Como paso previo al diseño de las superposiciones de los modelos de variograma anisotrópico se deben insertar los variogramas experimentales en las direcciones principales de anisotropía, dadas por los ejes X'' , Y'' , Z'' , como se describe en detalle en la sección 5.8.3. La interfaz permite el diseño simultáneo de los modelos de superposición para los variogramas directos, cruzados y secundarios con el fin de que el usuario pueda, si se presenta el caso, comprobar si se cumplen las condiciones necesarias para que el modelo de corregerionalización lineal construido sea utilizable en el procedimiento de estimación. Esto último implica que la matriz de covarianzas cruzada del modelo sea positiva semidefinida, como lo describe detalladamente Goovaerts (1997). En la versión actual de ESTELLUS no se ha implementado un algoritmo que efectúe automáticamente esta comprobación, con lo cual el usuario deberá realizarla por sus propios medios.

De esta forma, el ingreso de los variogramas experimentales se realiza mediante el menú **Archivo**, seleccionando el ítem correspondiente, de acuerdo a si se trabaja con el modelo de superposición para el variograma de la variable principal, la variable secundaria o el variograma cruzado (localizar cuadro amarillo en la figura 7.24). Por ejemplo, para insertar el variograma direccional correspondiente al eje de anisotropía X'' para la variable principal se pulsa sobre el ítem **Abrir variable principal, Variograma experimental dirección 1...**, de forma similar para los variogramas cruzados o secundarios. La convención utilizada en ESTELLUS para las direcciones de los ejes principales de anisotropía es: (i) dirección 1 corresponde a la dirección dada por el eje X'' , (ii) dirección 2 corresponde a la dirección dada por el eje Y'' y (iii) dirección 3 corresponde a la dirección dada por el eje Z'' . Se remite al usuario a la sección 5.8.3 para visualizar la relación de estas direcciones con los ángulos acimut, dirección e inmersión que las definen unívocamente.

Una vez se haya culminado el proceso de modelación, la superposición seleccionada actualmente puede almacenarse en disco. Para ello, en el menú **Archivo** se escoge la opción **Guardar corregerionalización...**, que da acceso a un selector de archivos donde el usuario elige la ubicación del directorio y el nombre del archivo de texto, con extensión ***.sup**. Al efectuar la operación de almacenamiento se produce un archivo de texto en formato especial donde se escriben de forma ordenadas los parámetros que componen cada uno de los modelos de variograma que forman parte de la superposición. El archivo ***.sup** puede editarse y visualizarse mediante cualquier procesador de texto. Una vez que se ingresa el modelo de variograma experimental en la dirección correspondiente se genera un gráfico en el panel en la pestaña correspondiente (**Principal**, **Cruzado** o **Secundario**) del panel **Ajuste de parámetros (1)**, como se explica en el siguiente apartado.

Por otro lado, en el menú **Opciones** se encuentra la posibilidad de habilitar el análisis de modelos lineales de corregerionalización tridimensional pulsando el ítem **Análisis**

tridimensional, con lo cual cambian las opciones de visualización y se activan en el menú **Archivo** los items asociados con la tercera dimensión (dirección 3).

Panel Ajuste de parámetros

El panel **Ajuste de parámetros**, denotado con el número (1) en la figura 7.24, permite al usuario ingresar los parámetros de los modelos de superposición lineal de variograma anisotrópicos principal, cruzado y secundario y visualizar los modelos diseñados en forma gráfica y de tablas. El panel esta formado por diferentes pestañas, nombradas como **Principal**, **Cruzada**, **Secundaria** y **Tablas**. En las tres primeras pestañas se ofrece al usuario las herramientas para el ingreso propiamente dicho de los parámetros. Dichas pestañas están formadas por:

- **Gráfico de variogramas en las direcciones principales de anisotropía:** en este gráfico se presentan de forma simultánea los variograma experimentales, ingresados desde archivo, y los variogramas correspondientes al modelo de superposición lineal en las direcciones dadas por los ejes X'' , Y'' , Z'' . Esto permite al usuario visualizar el comportamiento en dichas direcciones del modelo ajustado y su relación con los variogramas experimentales.
- **Filtro de dominio y rango:** son campos que le permiten al usuario realizar un filtro de la información visible en el gráfico de variogramas, estableciendo los valores máximos visibles del dominio (distancia) y del rango (valor del variograma) en el gráfico.
- **Panel Parámetros:** en los campos localizados en este panel se hace el ingreso de parámetros propiamente dicho. Inicialmente, debe establecerse el número de modelos de variograma anisotrópicos que se desea incluir en el análisis para que se activen automáticamente los restantes campos. Seguidamente, se selecciona la estructura a la cual se le desea modificar los parámetros y se introducen los valores deseados de estos en los campos correspondientes (pepita, contribución, rango, tipo de estructura, acimut, dirección e inmersión). El procedimiento se repite para cada modelo incluido, hasta lograr el diseño deseado por el usuario.

El ingreso de la información de parámetros se hace para cada tipo de superposición de modelos de variograma, es decir, principal, cruzado y/o secundario en la pestaña correspondiente. La visualización de los modelos construidos para cada uno de ellos se realiza por separado.

Panel Características del raster

En ESTELLUS, la visualización de la superposición lineal de modelos anisotrópicos de variograma se presenta mediante campos raster. Previo a la estimación de los valores asociados a cada celda del raster que representa el modelo construido, deben definirse

sus características básicas, coordenadas de la esquina inferior izquierda, tamaño de celdas y número de filas, columnas y capas, según lo descrito en la sección 4.4. El panel **Características del raster**, denotado con el número (2) en la figura 7.24 ofrece al usuario los campos de texto necesarios para el ingreso de dichos valores y, así, lograr la definición del dominio espacial de la superficie teórica del variograma que se va a desplegar en el panel de visualización (4).

Panel Seleccionar regionalización

El panel **Seleccionar regionalización**, denotado con el número (3) en la figura 7.24, contiene la opciones finales necesarias para visualizar los campos que representan las superficies de variogramas para las superposiciones lineales de modelos anisotrópicos. Habiendo ingresado los parámetros requeridos por el cálculo, debe elegirse de la caja de selección el variograma (principal, cruzado o secundario) que se desea visualizar, para, finalmente, presionar el botón **Aceptar** con el fin de visualizar el despliegue gráfico del modelo construido. Si se desea abortar el procedimiento y cerrar el cuadro de diálogo, se pulsa el botón **Cancelar**.

Panel de visualización

En el panel de visualización, denotado con el número (4) en la figura 7.24, se muestra el campo raster que representa la superposición de modelos anisotrópicos de variograma construido por el usuario. En este caso las utilidades de despliegue son iguales a las presentadas para la interfaz principal de ESTELLUS, por lo cual se recomienda al lector ir a la sección 7.2, donde se describe con detalle las utilidades implementadas para los fines de visualización de campos raster bidimensionales y tridimensionales.

7.8. Herramientas para la estimación

En esta sección se describen con detalle las interfaces en ESTELLUS para posibilitar al usuario el uso de los diferentes algoritmos de interpolación implementados. Las interfaces diseñadas se encuentran en el panel de algoritmos, (6) en la figura 7.2, de la interfaz principal.

También, se han implementado algoritmos de búsqueda de localizaciones por proximidad y por bloques, los cuales cobran importancia preponderante en problemas de estimación donde se tiene gran cantidad de información y se desea optimizar el tiempo de ejecución del algoritmo; junto con herramientas para realizar la validación cruzada de los diferentes estimadores y desplegar los resultados en forma de gráficas y tablas. En esta sección se presentan primero algunas opciones comunes a todos los algoritmos de estimación, para posteriormente describir las opciones particulares para cada tipo de estimador.

7.8.1. Acceso a los diferentes algoritmos de estimación

En la sección 7.2 se describieron las componentes de la interfaz principal de ESTELLUS. De ellos algunos estaban especialmente diseñados para proporcionar al usuario las herramientas necesarias para proceder a realizar la estimación de campos aleatorios. Haciendo referencia a las figuras 7.2 y 7.3, en ellas se observan los paneles numerados como (5), (6) y (7). El panel (5) permite al usuario la selección del algoritmo de estimación que desee utilizar. Hecha la selección, la ventana correspondiente se despliega en el panel (6), en la figura 7.2, ofreciendo al usuario las diferentes herramientas computacionales para ejecutar la estimación. Finalmente, en el panel (7) se despliegan las opciones para ejecutar el algoritmo, la validación cruzada y una barra de progreso que informa al usuario del avance de la estimación. En las siguientes secciones se describen las opciones de los paneles diseñados para la ejecución de los algoritmos de estimación.

7.8.2. Interfaz para el krigado ordinario y simple

En la figura 7.25 se presentan las interfaces diseñadas para ejecutar los algoritmos de krigado ordinario (*KO*) y krigado simple (*KS*). Proporcionan al usuario las herramientas para ingresar los diferentes parámetros necesarios para la ejecución de ambos algoritmos. Lo anterior equivale a que, una vez se han ingresado los requerimientos de la interfaz, se construye un objeto de las clases correspondientes (*KrigingOrd2D*, *KrigingOrd3D*, *KrigingSim2D* y *KrigingSim3D*) y es posible realizar la ejecución de los métodos *InterpolarMalla* presentados en el capítulo 6. Las ventanas están compuestas de cuatro pestañas de acceso: *Datos*, *Variograma*, *Mapa*, y *Búsqueda*, las cuales contienen las opciones que se describen a continuación.

La pestaña Datos

La pestaña *Datos* esta formada por tres paneles, que han sido numerados como (1), (2) y (9), como puede observarse en la figura 7.25. El panel *Ingreso de datos* (1) permite la entrada de la información de coordenadas y los valores del atributo asociado desde un archivo o el uso de la lista de observaciones activa en la lista de objetos de la interfaz principal de ESTELLUS. Al momento de seleccionar alguna de las dos opciones, en los campos *Nombre variable*, *Sistema de coordenadas* y *Ruta archivo fuente* se escriben la variable que se va a estimar, el sistema de georeferenciación de las coordenadas de las localizaciones disponibles y, en caso de que se haya seleccionado desde archivo, la ruta de localización del archivo de datos, en formato nativo **.txt* de ESTELLUS, como se describió en la sección 7.2.1. Empleando la caja de selección *Georeferenciación plana* el usuario selecciona el sistema de georeferenciación en el cual desea calcular las distancias euclídeas, en el caso en que el sistema de coordenadas de las localizaciones de las entidades puntuales sea geográfica. Para el panel asociado con *KS* se incluye el campo para la definición de la media regional de la variable a estimar en el dominio de trabajo. En éste puede escribirse el valor de la media estimada a priori por el usuario o, si él lo desea, puede usar

por defecto la media estimada de la muestra de observaciones disponibles, seleccionando la caja **Estimar media**.

La pestaña **Datos** también ofrece al usuario la posibilidad de visualizar los estadísticos descriptivos de la muestra de observaciones que se está utilizando como información en la estimación de un campo, los cuales se despliegan en la tabla **Estadísticos generales**, como se muestra en el panel denotado mediante el número (2), en la figura 7.25. Antes de proceder con las estimaciones, el usuario debe adoptar una convención para las observaciones faltantes que se presenten en el archivo insertado o en la lista de observaciones elegida desde la interfaz principal de ESTELLUS. Dicho valor faltante se escribe en el campo de texto **Valor de faltante** y entonces se puede proceder a realizar un análisis descriptivo de los datos haciendo click en **Estadísticos**. Al hacerlo, los descriptores de la muestra estadística que se despliegan en la tabla **Estadísticos generales** son: mínimo, máximo, rango, media, mediana, varianza, desviación estándar, coeficiente de variación coeficiente de asimetría y número de faltantes en la muestra.

Las pestañas Variograma, Mapa y Búsqueda

En la figura 7.26 se presentan las opciones consignadas en las pestañas **Variograma**, **Mapa** y **Búsqueda** para los algoritmos de *KO* y *KS*. El panel **Variograma** ofrece al usuario la posibilidad de ingresar la información del modelo espacial de continuidad, lo cual se puede hacer de dos maneras: (i) seleccionando un archivo ***.sup** que contiene los valores de los parámetros de los diferentes modelos de variograma; para lo cual se accede desde el panel denotado con el número (3) en la figura 7.26, haciendo click sobre el botón asociado y usando un selector de archivo; y (ii) ingresando la información del modelo de variograma manualmente desde el panel **Parámetros** denotado con el número (4) en la figura 7.26.

La entrada manual de variogramas se realiza definiendo inicialmente el número de estructuras que van a componer el modelo de continuidad espacial, sin tener en cuenta el efecto pepita, en el campo **Número de estructuras**. Con ello se activan los campos de texto que permiten la entrada de los valores de los parámetros y la lista de selección ubicada en la parte superior del panel. Seguidamente, de la lista se debe seleccionar el modelo componente al cual se le desea modificar los valores de los parámetros y proceder a llenar los campos correspondientes: **Contribución estructura**, **Tipo de estructura**, **Alcances** para las diferentes direcciones de anisotropía y los ángulos que definen la orientación de los ejes principales de anisotropía, **Acimut**, **Dirección** e **Inmersión**, de acuerdo con las convenciones presentadas en la sección 5.8.3. Para obtener el conjunto de parámetros, este procedimiento debe repetirse para cada estructura componente de la superposición lineal de modelos. Los cambios realizados por el usuario se consignan en la tabla asociada a la lista, donde pueden observarse los valores de los parámetros de todos los modelos de variograma involucrados en la estimación. También, en la pestaña **Variograma**, en especial en el panel denotado por el número (5), el usuario puede definir si desea plantear el sistema de krigeado, dados en las ecuaciones 6.7 y 6.10, en términos del variograma o las covarianzas y si se desea efectuar la estimación usando un algoritmo de búsqueda.

La pestaña **Mapa** permite al usuario definir las propiedades del dominio espacial y los metadatos del campo raster que van a ser utilizados durante el procedimiento. En el panel **Geometría del raster**, denotado con el número (6) en la figura 7.26, se introducen los valores que forman los metadatos del campos raster: coordenadas de la esquina inferior izquierda, tamaño de celdas, y número de filas, columnas y capas, además del sistema de georeferenciación del campo estimado. Finalmente, si el usuario ha seleccionado la opción de realizar un algoritmo de búsqueda en el proceso de estimación deberá completar las entradas definidas en la pestaña **Búsqueda**, la cual está compuesta por los paneles **Elipsoide de búsqueda** y **Opciones de búsqueda**, denotados con los números (7) y (8) como se observa en la figura 7.26. Como se describió en la sección 6.7, en ESTELLUS se ha implementado dos tipos de algoritmos de búsqueda: por áreas vecinas y por bloques. La definición de parámetros de cada técnica es la siguiente:

- **Búsqueda por áreas vecinas:** se basa principalmente en la detección de las localizaciones puntuales con información que se encuentran dentro de un elipsoide de búsqueda definido por el tamaño de sus semiejes y por los ángulos de orientación de sus ejes principales. Además, puede incluirse un factor de dilatación o contracción para que el algoritmo aumente o disminuya la zona de búsqueda de acuerdo con la definición del número mínimo o máximo de localizaciones identificadas dentro de la zona de búsqueda. Los valores para estos parámetros se definen en el panel (8) en los campos **Semieje mayor**, **Semieje intermedio**, **Semieje menor**, **Acimut de búsqueda**, **Dirección de búsqueda**, **Inmersión de búsqueda**, **Factor de ampliación** y **Factor de reducción**. De forma similar, en el panel (7) se definen los límites de localizaciones dentro del elipsoide de búsqueda en los campos **Nro. máximo de puntos** y **Nro. mínimo de puntos**.
- **Búsqueda por bloques:** se basa principalmente en una clasificación por zonas de las observaciones disponibles. Para ello deben definirse las características geométricas de las regiones de clasificación. La idea es superponer un raster mucho mas grueso que el original, por lo cual deben definirse las coordenadas máximas y mínimas sobre las cuales se desea hacer la clasificación en los campos **Mínimo X Reclasificación**, **Mínimo Y Reclasificación** y **Mínimo H Reclasificación**, para las coordenadas mínimas y **Máximo X Reclasificación**, **Máximo Y Reclasificación** y **Máximo H Reclasificación**, para las coordenadas máximas, como se presenta en la figura 7.26, en los paneles (7) y (8). El numero de sectores se ha definido previamente como 4 por cada eje coordenado. Ahora bien si se desea que el algoritmo de clasificación sea más eficiente al momento de estimar se selecciona la búsqueda por espiral, con lo cual debe definirse el número máximo de puntos para interpolar.

El panel de ejecución

Cuando el usuario ha culminado la entrada de los parámetros del algoritmo (sea *KO* o *KS*) puede presionar el botón **Aceptar**, que procede a ejecutarlo. Éste se localiza en el

panel denotado por el número (9) en la figura 7.26. Una vez se ha iniciado la ejecución de la estimación, su avance puede leerse de la barra de progreso adyacente al botón **Aceptar**. Cuando culmina el proceso de estimación aparece un cuadro de diálogo que informa que la estimación ha sido exitosa y el campo resultado debe guardarse en archivo para ser posteriormente insertado en la lista correspondiente a los objetos de la interfaz principal de ESTELLUS.

7.8.3. Interfaz para la estimación mediante los métodos deterministas

Las interfaces creadas para los métodos deterministas son similares a las creadas para el *KO* y el *KS*, pero no presentan la pestaña diseñada para el ingreso de los parámetros de los modelos espaciales diseñados, paneles (3), (4) y (5) en la figura 7.26. En el caso del panel contenedor para el método del inverso cuadrático elevado a un exponente se ha adicionado al panel (1), **Ingreso de datos**, de la figura 7.25 el campo de texto **Exponente** que permite al usuario ingresar el exponente deseado.

7.8.4. Interfaz para el krigeado con una deriva externa

En la figura 7.27 se presenta la interfaz de usuario diseñada para la ejecución de la estimación espacial usando krigeado con una deriva externa (*KDE*) incluida en las herramientas ESTELLUS. De forma similar a las implementadas para *KO* y *KS*, el panel contenedor diseñado para *KDE* está compuesto por las pestañas **Datos**, **Variograma**, **Mapa** y **Búsqueda**. De ellas las dos últimas tienen funciones idénticas a las presentadas para *KO* y *KS*, en la sección 7.8.2, mientras que la pestaña **Datos** para *KDE* requiere la inclusión de opciones adicionales para permitir el ingreso de parámetros propios del algoritmo. Dichas opciones se describen a continuación.

Las pestaña Datos

La pestaña **Datos** para *KDE* presenta las mismas opciones que la análogo para las técnicas *KO* y *KS*. Adicionalmente, se incluye la posibilidad de ingresar el campo de información secundaria en el panel **Ingreso de datos**. Para ello el usuario tiene dos opciones: (i) ingresar el campo desde un archivo en formato nativo de ESTELLUS, con extensiones ***.metaGEO** o ***.meta3GEO**, para lo cual debe pulsar el botón **Mapa de derivas** y elegir la ubicación mediante un selector de archivos asociado a éste; y (ii) seleccionar el campo raster activo actualmente en la interfaz principal de ESTELLUS. La interfaz detecta automáticamente la dimensión espacial del procedimiento de estimación en base al tamaño del vector de coordenadas de la lista de observaciones; si su tamaño es dos, la interfaz sólo admite el ingreso de campos raster tipo bidimensionales, con extensión ***.metaGEO**. Análogamente, si el tamaño de la matriz de coordenadas es tres, lo anterior se aplica para los archivos con extensión ***.meta3GEO**.

El panel de ejecución

Cuando el usuario ha culminado el ingreso de los parámetros del algoritmo, puede presionar el botón **Aceptar** para ejecutar el algoritmo seleccionado, el cual está localizado en el panel denotado por el número (9), como se observa en la figura 7.27. El funcionamiento del panel es idéntico al presentado para los algoritmos de *KO* y *KS*.

7.8.5. Interfaz para el cokrigado ordinario y ordinario estandarizado

Como puede observarse en la figura 7.28, las interfaces para los algoritmos de cokrigado ordinario *CKO* y cokrigado ordinario estandarizado *CKOE*, están compuestas por cuatro pestañas: **Datos**, **Variograma**, **Mapa** y **Búsqueda**. El funcionamiento de estas pestañas es similar a sus análogas en las interfaces para los algoritmos de *KO* y *KS*. A pesar de ello, debido a la inclusión de las variables auxiliares y los modelos de variación espacial relacionados (cruzado y secundario), estas ventanas ofrecen una mayor cantidad de opciones. Los apartados siguientes describen las funciones de cada pestaña.

La pestaña Datos

La pestaña **Datos** correspondiente a los algoritmos de *CKO* y *CKOE* presenta las mismas opciones que sus análogas para *KO* y *KS*, con la excepción de que se incluye un panel adicional para ingresar las coordenadas y los valores del atributo secundario, como se observa en la figura 7.29 en el panel enmarcado por un cuadro rojo con el número (1). Estos paneles de **Variables secundarias** ofrecen la posibilidad de ingresar la información desde un archivo de texto, que debe tener el formato nativo de ESTELLUS, o desde las listas de observaciones cargadas desde la interfaz principal de la aplicación. Si se trata del segundo caso, la lista de observaciones asociada con la variable secundaria debe escogerse en un cuadro de diálogo externo para que el usuario pueda proceder a realizar el ingreso de los valores para los parámetros. Por otro lado, si se trata del algoritmo de *CKOE* uno de los parámetros a tener en cuenta es el valor de las medias de ambas variables involucradas en la estimación. En la interfaz diseñada el usuario puede ingresar un valor determinado a priori para ellas o puede estimarla a partir de las observaciones disponibles.

La tabla de información acerca de los estadísticos descriptivos también es diferente con respecto a las interfaces de estimación previamente descritas. Como puede verse en la figura 7.29, en el panel **Estadísticos generales**, denotado por el número (2), se tienen dos pestañas en las cuales puede consultarse los estadísticos descriptivos de las muestras de observaciones de las variables principal y secundaria.

La pestaña Variograma

En la pestaña **Variograma** el usuario puede ingresar los modelos de superposición lineal de variogramas anisotrópicos involucrados en la estimación. Como se presenta en la figura 7.28, esta pestaña, a su vez, está compuesta de tres pestañas nombradas **Principal**,

Cruzado y Secundario; en cada una de la cuales el usuario debe ingresar un modelo de variación diseñado para la variables principal y secundaria y el variograma cruzado, respectivamente. Todas las pestañas tienen exactamente la misma configuración y ofrecen al usuario la posibilidad de ingresar los parámetros de cada modelo de variograma de dos formas: (i) desde archivo (un archivo de texto con extensión `*.sup`), en el panel denotado por el número (3) (figura 7.28), y (ii) manualmente en los campos de texto localizados en el panel **Parámetros**, que se denota como (4) en la figura 7.28. El funcionamiento operativo de cada pestaña es idéntico al descrito para la pestaña **Variograma** en las interfaces para *KO*, *KS* y *KDE*.

La pestaña Mapa

La pestaña Mapa es idéntica a las presentadas para *KO*, *KS* y *KDE*. En ella se definen las características geométricas del campo raster de salida.

La pestaña Búsqueda

La pestaña Búsqueda permite al usuario definir los parámetros de búsqueda tanto para la variable principal como para la variable secundaria. Así, en las interfaces para la ejecución de *CKO* y *CKOE* se han implementado búsquedas independientes para las entidades puntuales pertenecientes a la variable principal y a la variable secundaria. Lo anterior se estableció considerando que, en la mayoría de los casos prácticos, existe mayor densidad de observaciones disponibles para una variable que para la otra. Cada pestaña está compuesta por los paneles **Opciones de región**, **Elipsoide de búsqueda** y **Opciones de búsqueda**, numerados respectivamente (7) y (8), como se puede ver en la figura 7.28, en los cuales están los campos de texto para ingresar los parámetros del algoritmo de búsqueda respectivo. El significado de cada parámetro puede consultarse en la sección 6.7. También el funcionamiento de los componente de estos paneles es idéntico al de sus análogos implementados para *KO* y *KS*.

El panel de ejecución

Cuando el usuario ha culminado el ingreso de los parámetro del algoritmo (sea *CKO* o *CKOE*), puede proceder a presionar el botón **Aceptar** que ejecuta el algoritmo. Nuevamente, éste se localiza en el panel denotado por el número (9) en la figura 7.28. El funcionamiento de este panel es idéntico al presentado para los algoritmos de *KO* y de *KS*.

7.8.6. Interfaz para la ejecución del cokrigado colocalizado bajo un modelo de Markov

En la figura 7.30 se presenta la interfaz de usuario diseñada para la ejecución de la estimación espacial usando cokrigado colocalizado bajo un modelo de Markov (*CKCM*).

Puede verse que está compuesta por cuatro pestañas denominadas: **Datos**, **Variograma**, **Mapa** y **Búsqueda**. Las funcionalidades, componentes e información de ingreso requerida por **Datos**, **Variograma**, y **Búsqueda** son idénticas a las presentadas para *KO*.

La pestaña Mapa

La pestaña **Mapa** está conformada por una barra de herramientas y un panel de información, numerados como (10) y (11), respectivamente, en la figura 7.30, junto con el panel de **Geometría del mapa**, el cual tiene una configuración idéntica que para las demás interfaces de los estimadores hasta aquí presentados. La barra de herramientas (10) ofrece al usuario las utilidades para ingresar y visualizar las características de la información secundaria y para activar su funcionamiento debe haberse ingresado información acerca de la variable principal. Los botones de herramientas que la componen son los siguientes:

-  Insertar nuevo campo de información secundaria: este botón se utiliza para insertar desde archivo un nuevo campo de información auxiliar que va a ser integrada al procedimiento de estimación. Al pulsar el botón se despliega un selector de archivo en el cual el usuario escoge la localización de la fuente de información, en formato nativo de ESTELLUS (con extensiones `*.metaGEO` y `*.meta3GEO`). Si la lectura es exitosa, el nombre del campo registrado en los datos aparece en la lista de variables insertadas, en el panel **Información de variables auxiliares**, (11) en la figura 7.30. Una vez el usuario ha ingresado todos los campos auxiliares deseados, el algoritmo revisa si éstos manifiestan un idéntico dominio espacial y, de no hacerlo, la ejecución de la estimación se aborta. La única forma de evitar lo anterior es ingresar campos con valores idénticos para las coordenadas de la esquina inferior, tamaños de celda y número de filas, columnas y capas.
-  Eliminar campo de información secundaria: al pulsar este botón el usuario elimina el campo de información auxiliar correspondiente al ítem seleccionado en la lista de variables insertadas en el panel **Información de variables auxiliares**, (11) en la figura 7.30.
-  Visualizar campo de variable auxiliar: esta función abre un visor del mapa en una ventana nueva en la cual se muestra el campo seleccionado en la lista de variables insertadas en el panel **Información de variables auxiliares** (11), en la figura 7.30.
-  Visualizar diagrama de dispersión: esta herramienta permite al usuario extraer un diagrama de dispersión con la información colocalizada de las variables principal y secundaria seleccionada en la lista de variables insertadas en el panel **Información de variables auxiliares** (11), en la figura 7.30.

El panel **Información de variables auxiliares** ofrece al usuario las opciones de manipulación de la información secundaria insertada y hace posible la ejecución de las funciones disponibles en la barra de herramientas. Cuando el usuario selecciona un ítem insertado en la lista de variables, los siguientes cuadros de texto se llenan con la información del campo:

- **Nombre:** el nombre del campo consignado en los metadatos.
- **Coordenadas:** el sistema de coordenadas asociado con el dominio espacial del campo.
- **Ruta fuente:** es una cadena de caracteres con la localización del directorio contenedor del archivo en el disco.
- **Varianza:** es la varianza de la variable secundaria estimada a partir de los valores del campo.
- **Correlación lineal:** es la correlación lineal estimada utilizando los valores colocalizados de las variables principal y secundaria.

El panel de ejecución

Cuando el usuario ha culminado el ingreso de los parámetros del algoritmo, puede proceder a presionar el botón **Aceptar** para ejecutarlo. Dicho botón se localiza en el panel denotado por el número (9) en la figura 7.30. El funcionamiento de este panel es idéntico al presentado en secciones anteriores para los diferentes algoritmos implementados.

7.8.7. Ejecución de la técnica de validación cruzada y visualización de sus resultados

Para efectuar la validación cruzada de una estimación en ESTELLUS, basta con activar la caja de selección **Realizar validación cruzada al terminar la estimación**, disponible en el panel de ejecución de todas las interfaces que hacen posible realizar las estimaciones con los algoritmos implementados, descritos a lo largo de este capítulo. Los resultados del procedimiento de validación cruzada se presentan utilizando la interfaz de despliegue para los resultados de análisis estadísticos bivariados como se presenta en la sección 7.5.2, en la figura 7.20.

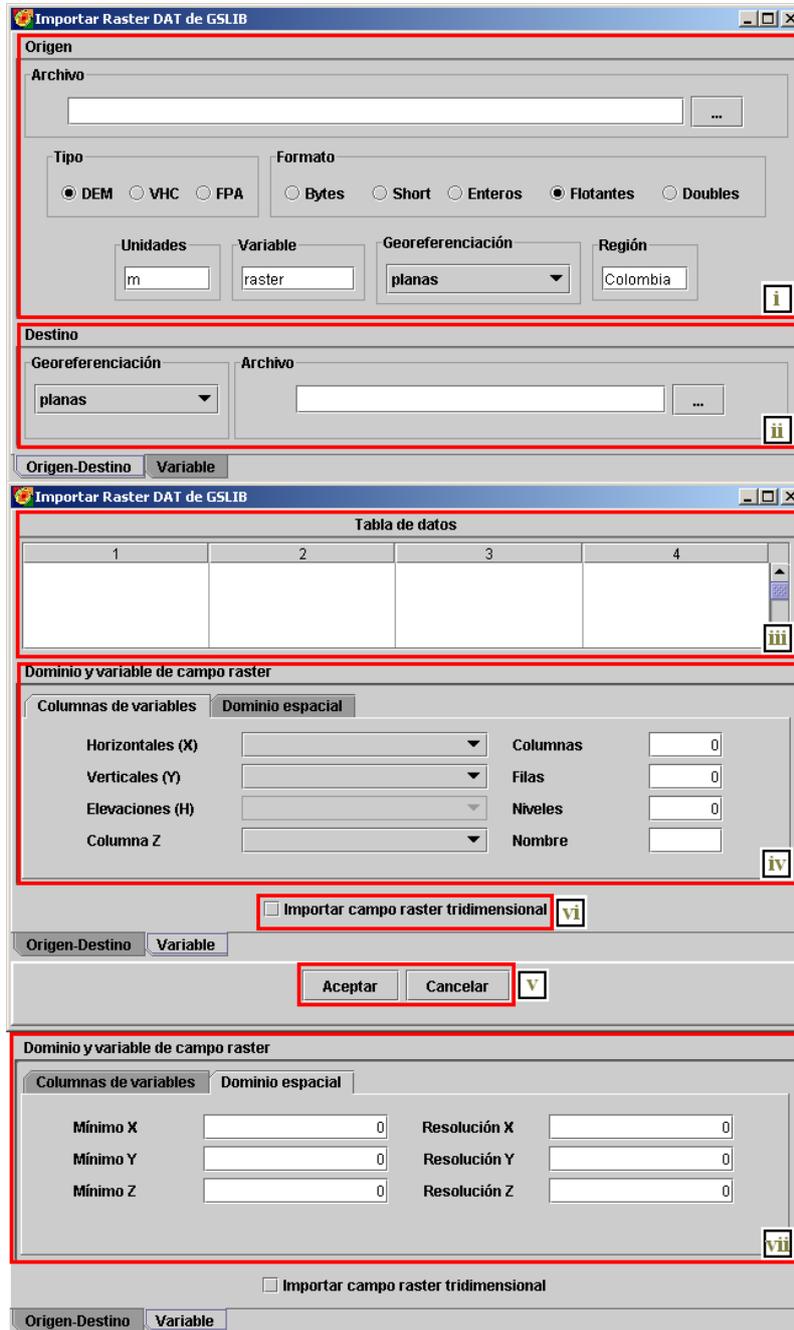


Figura 7.14. Cuadro de diálogo para importar información raster desde archivos en formato GEO-EAS.

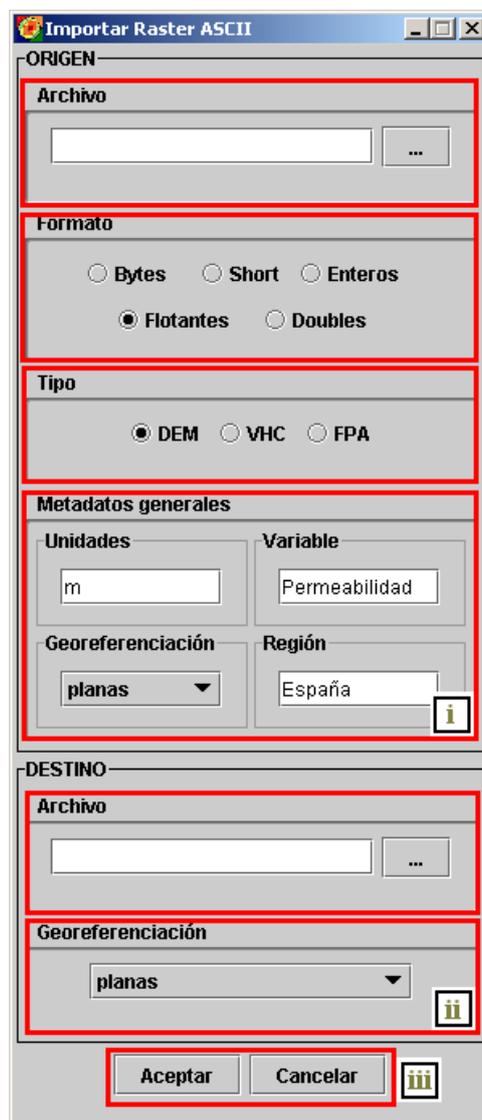


Figura 7.15. Interfaz implementada para importar archivos raster en formato ascii de ESRI.

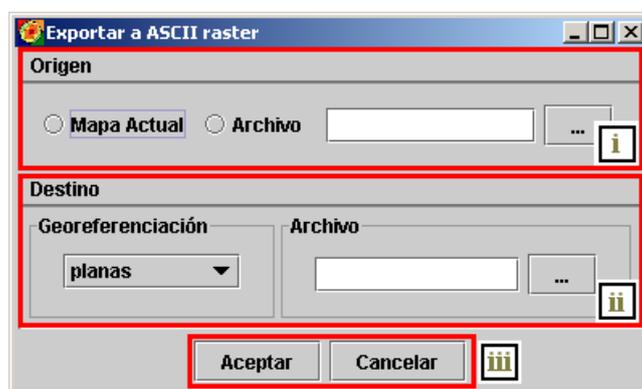


Figura 7.16. Interfaz implementada para exportar archivos raster en formato nativo de ESTEL-LUS a formato ascii raster de ESRI.

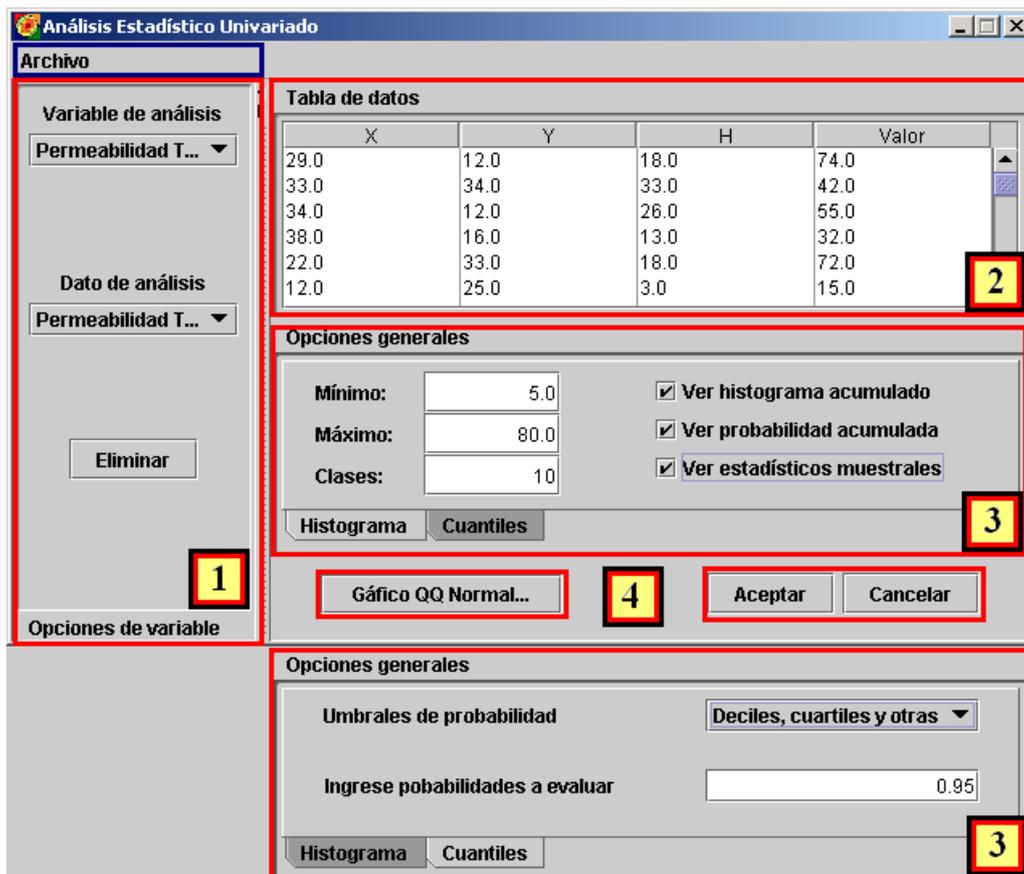


Figura 7.17. Interfaz implementada que realiza el análisis estadístico descriptivo univariado en ESTELLUS.

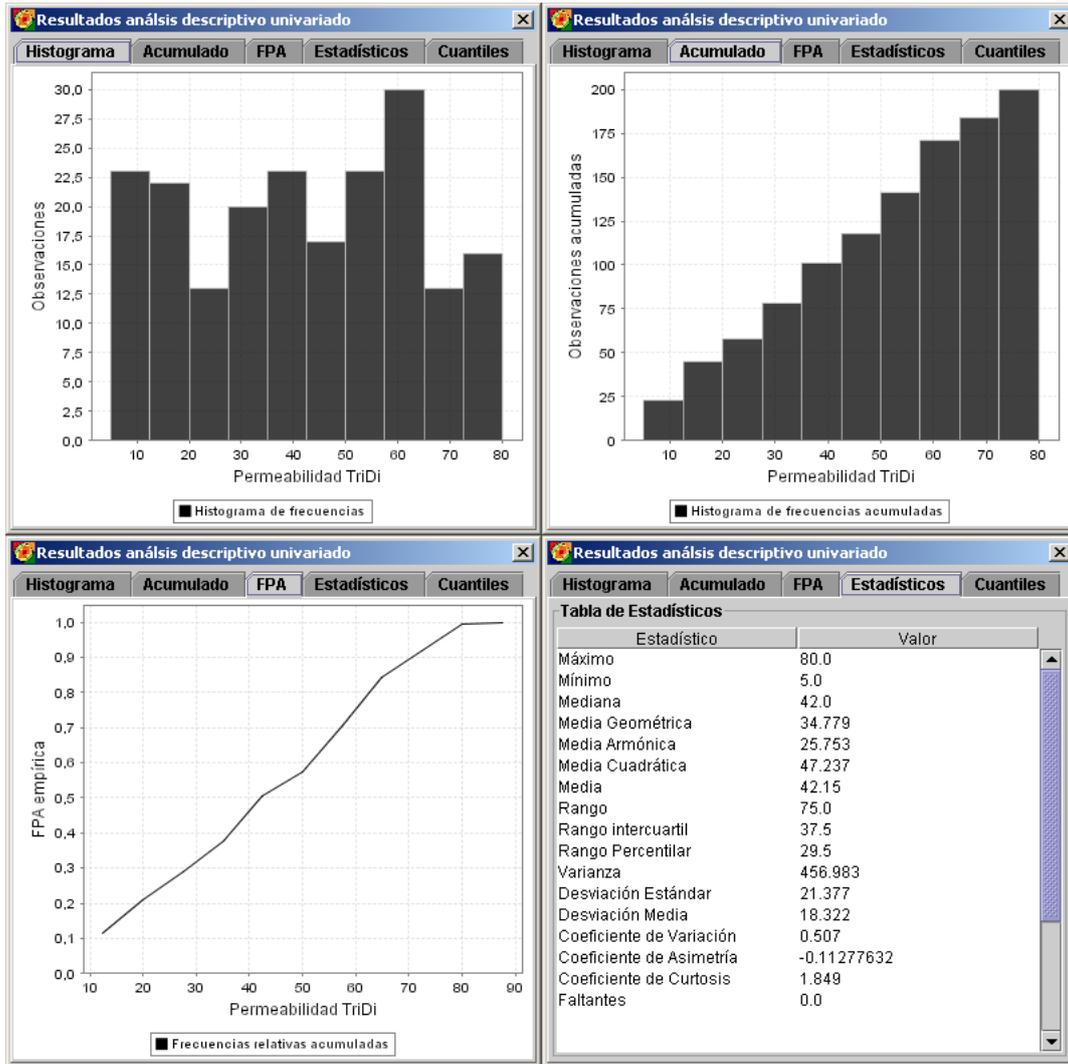


Figura 7.18. Gráficos y tablas resultados de realizar un análisis descriptivo univariado para una en ESTELLUS.

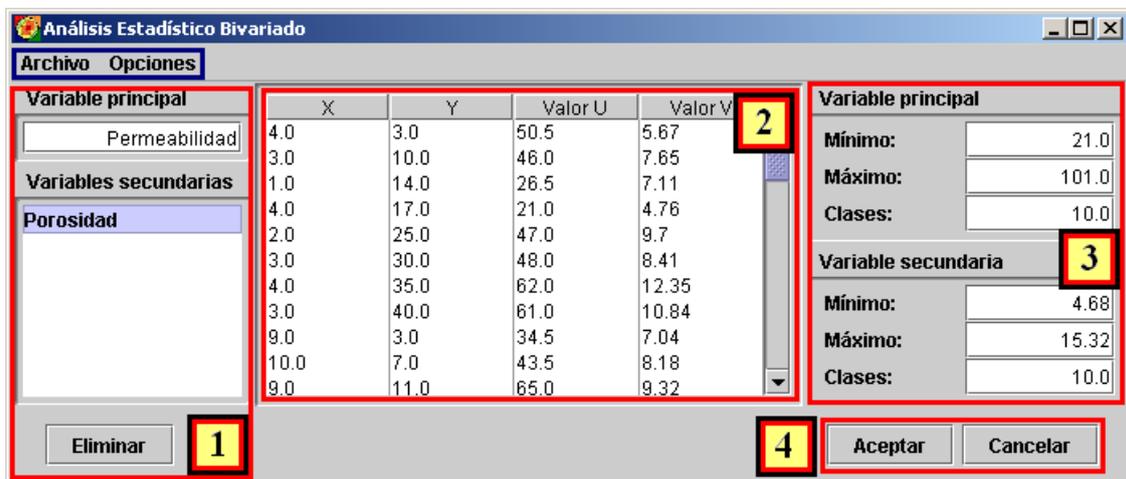


Figura 7.19. Interfaz implementada que realiza el análisis estadístico descriptivo bivariado en ESTELLUS.

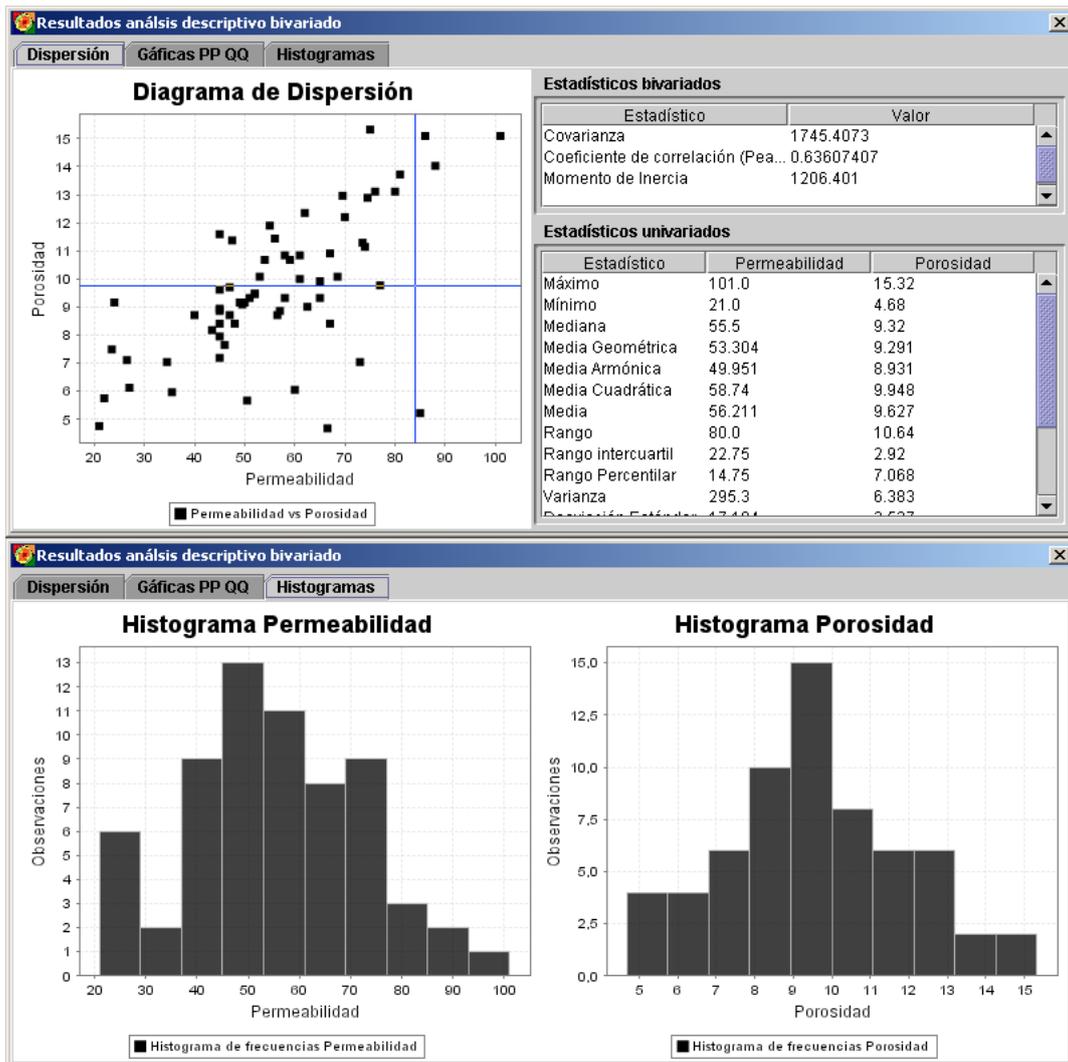


Figura 7.20. Gráficos y tablas resultados de realizar un análisis descriptivo bivariado para una en ESTELLUS

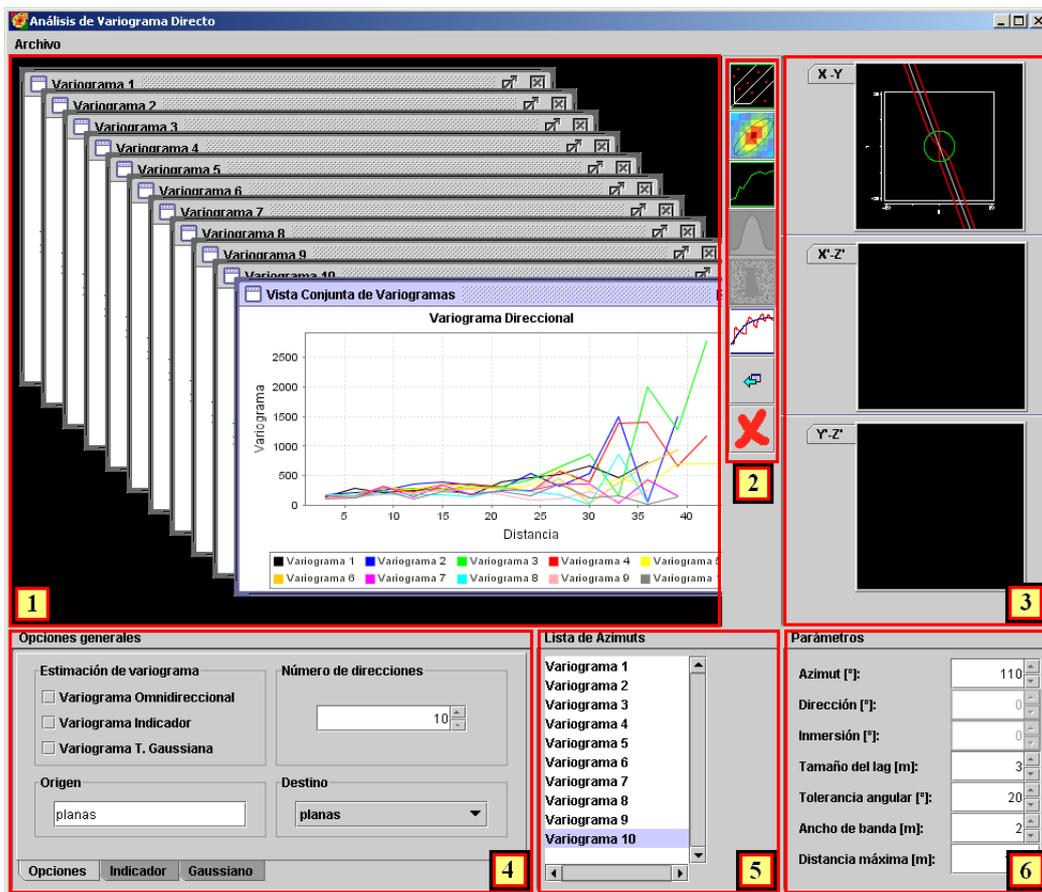


Figura 7.21. Interfaz diseñada en ESTELLUS para el cálculo de variogramas directos.

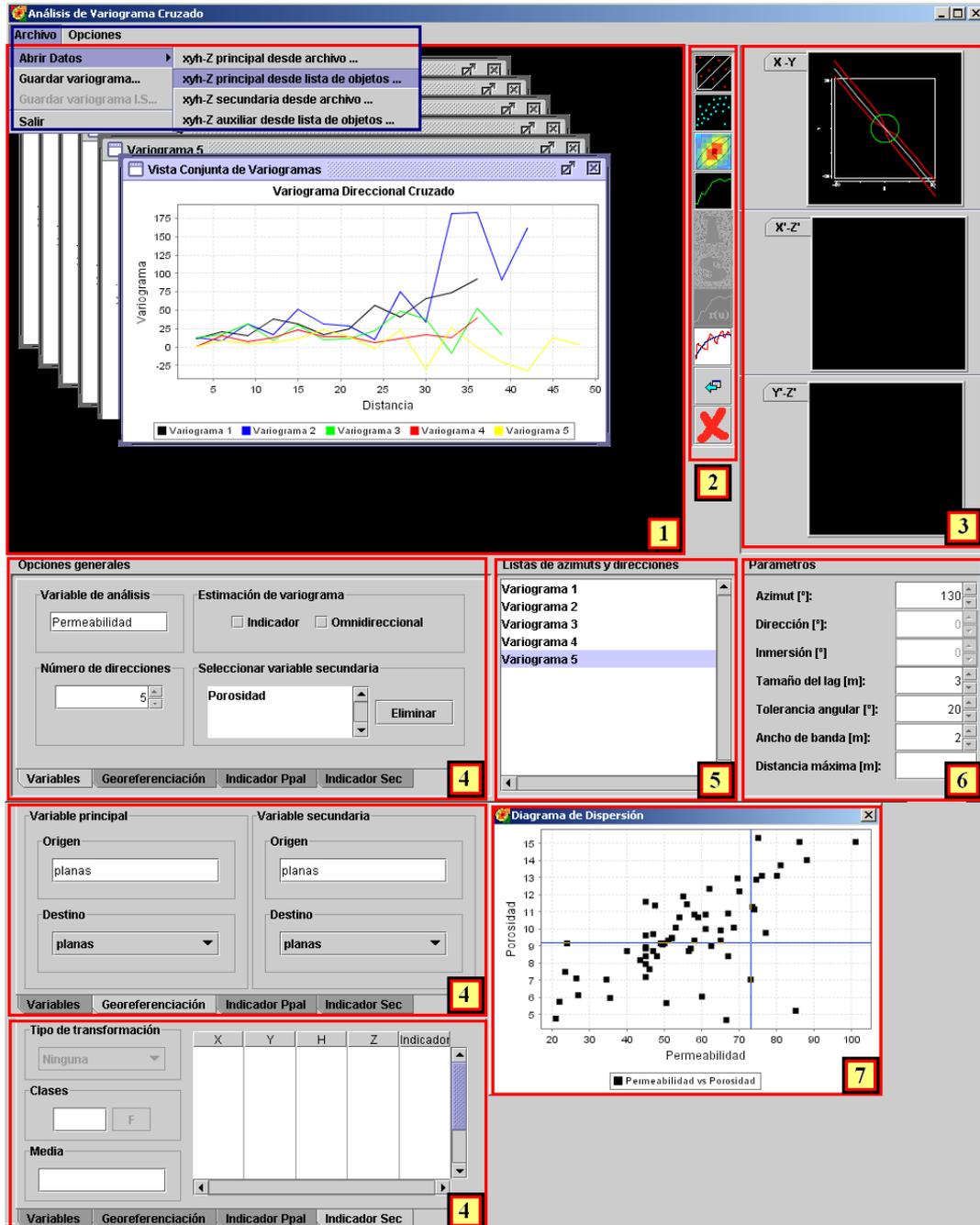


Figura 7.22. Interfaz diseñada en ESTELLUS para el cálculo de los variogramas cruzados.

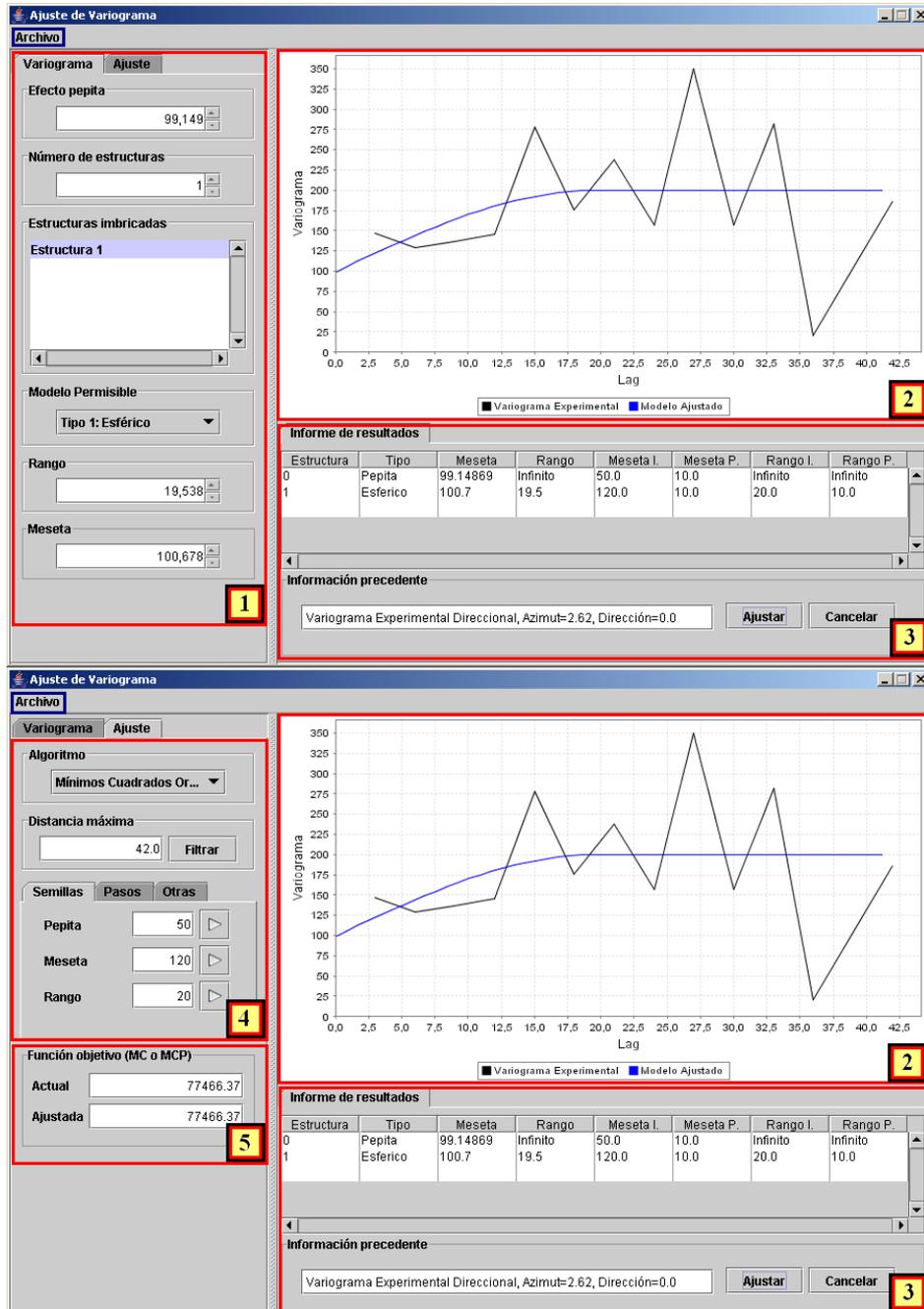


Figura 7.23. Interfaz diseñada en ESTELLUS para el ajuste de modelos a un variograma experimental.

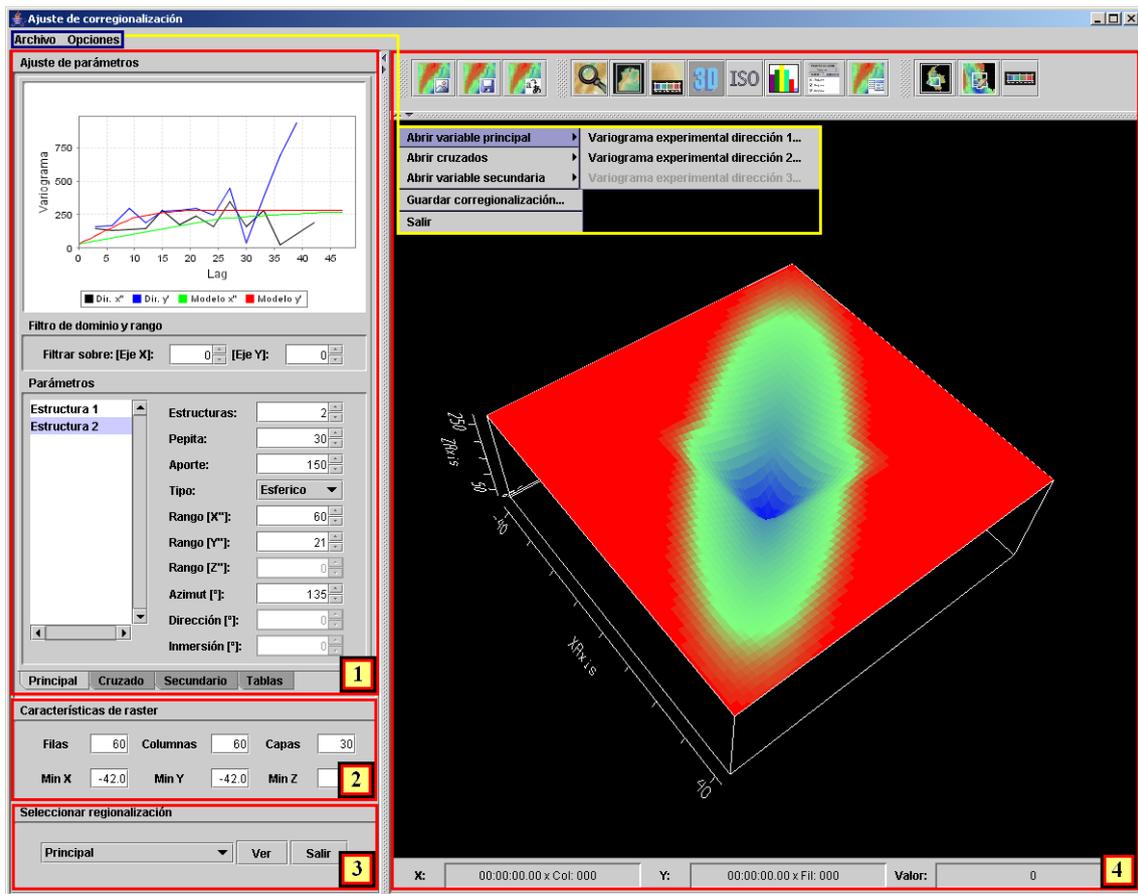


Figura 7.24. Interfaz diseñada en ESTELLUS para la creación de superposición lineal de modelos de variograma anisotrópicos.

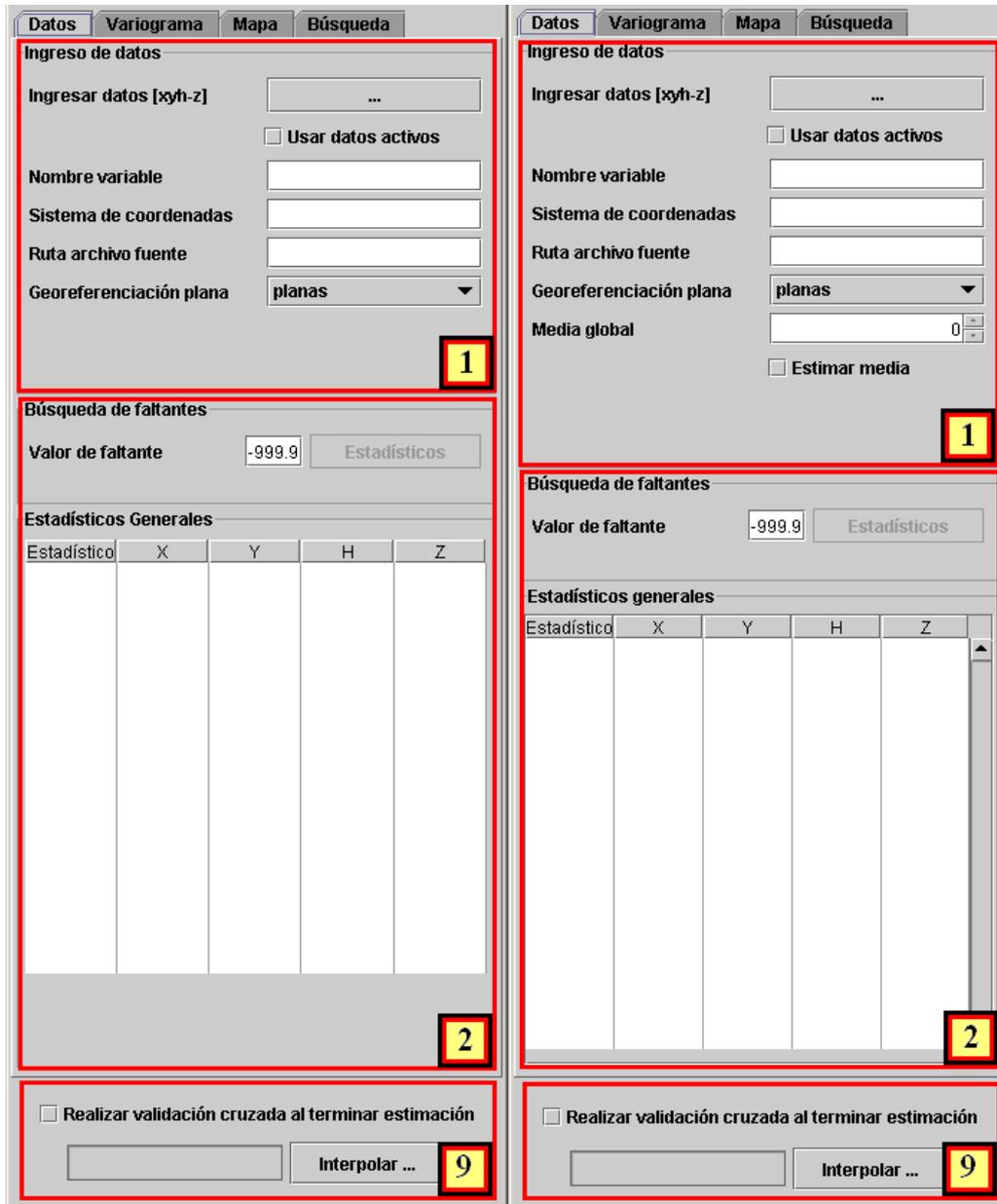


Figura 7.25. Vista general del panel diseñado para la ejecución de los algoritmos de krigado ordinario y simple.

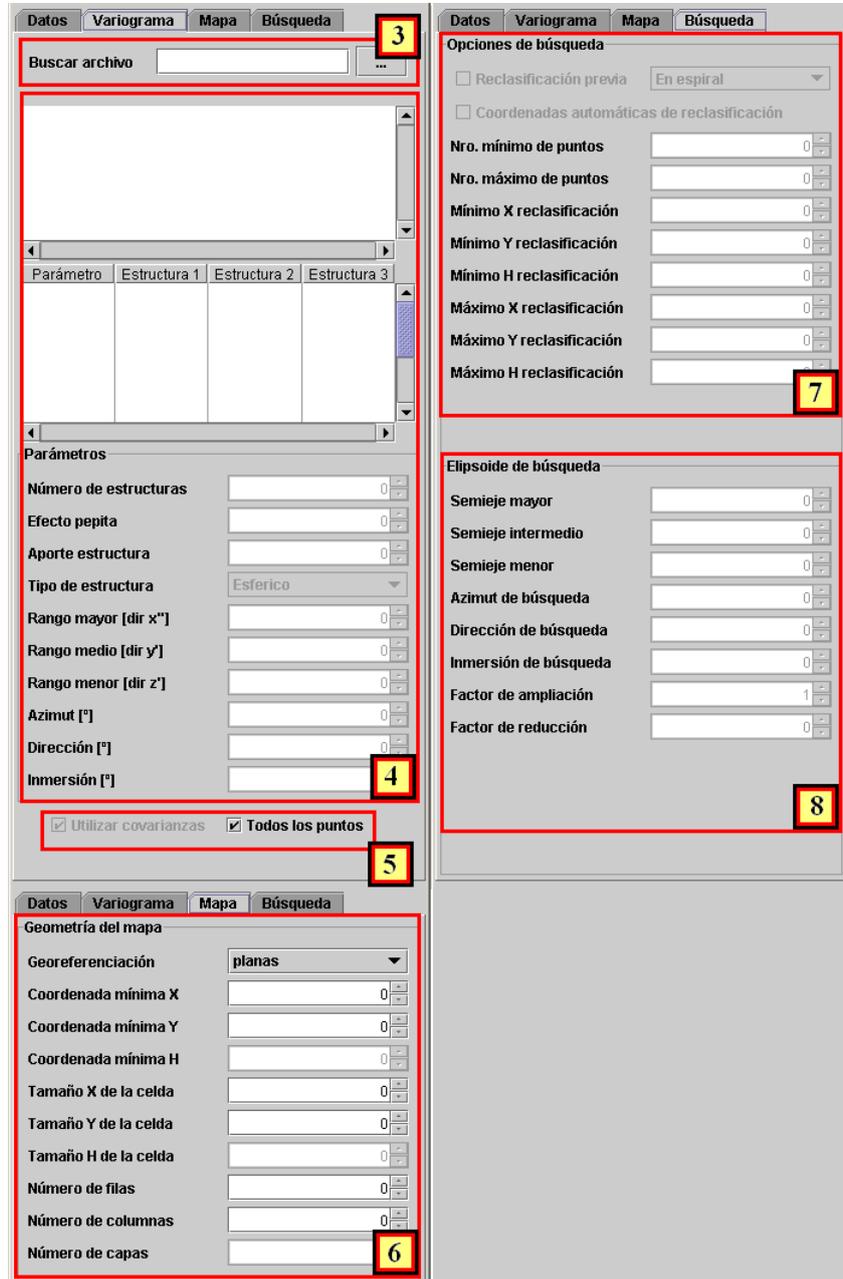


Figura 7.26. Panel diseñado en ESTELLUS para la entrada de los modelos de variograma, la información del dominio espacial del campo raster de salida y los parámetros del algoritmo de búsqueda para KO y KS

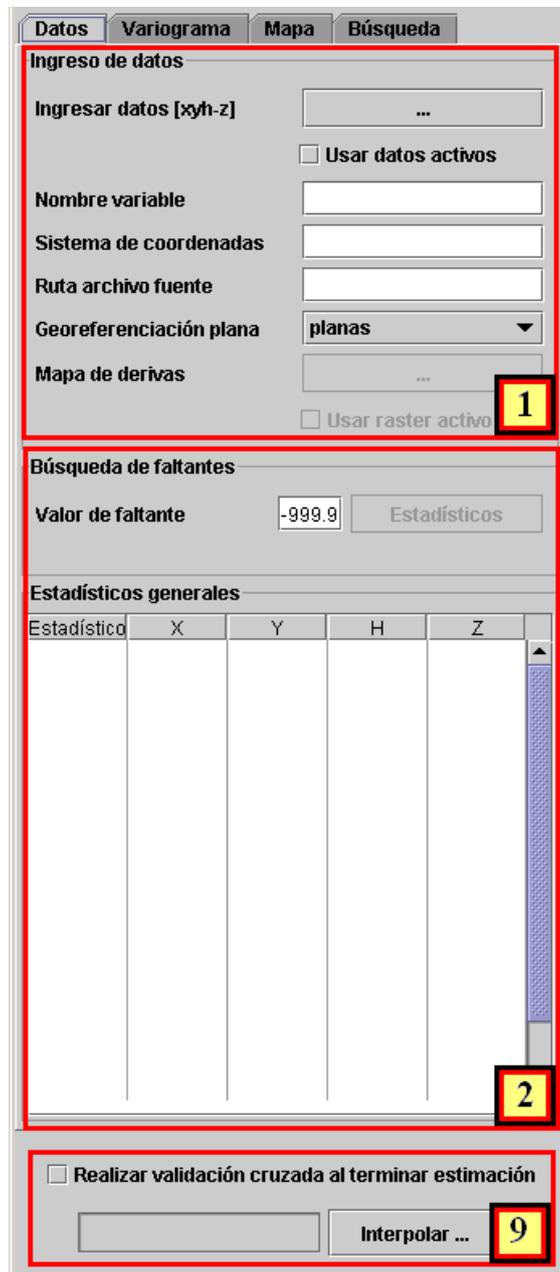


Figura 7.27. Vista general de los panel contenedores diseñados para la ejecución del algoritmos de estimación krigeado con deriva externa *KDE*.



Figura 7.28. Interfaces para la entrada de los modelos de continuidad espacial, dominio de salida y parámetros de búsqueda para el *CKO* y *CKOE*.

The image shows two side-by-side screenshots of a software interface for data processing. Each panel has a menu bar with 'Datos', 'Variograma', 'Mapa', and 'Búsqueda'. The left panel is divided into sections for 'Variable Principal' and 'Variable Secundaria', each with fields for 'Ingresar datos [xyz-h]', 'Nombre variable', 'Sistema de coordenadas', 'Ruta archivo fuente', 'Georeferenciación plana', and 'Media global'. The right panel has a similar structure but includes a 'Búsqueda de faltantes' section with a 'Valor de faltante' field and 'Estadísticos' button. Below this is an 'Estadísticos Generales' table with columns for 'Principal', 'Secundaria', 'Estadísti...', 'X', 'Y', 'H', and 'Z'. At the bottom of each panel is a checkbox for 'Realizar validación cruzada al terminar estimación' and an 'Interpolar ...' button. Red boxes outline the main input areas, and yellow boxes with numbers 1, 2, and 9 highlight specific elements: '1' is on the 'Estimar media' checkbox, '2' is on the 'Interpolar ...' button, and '9' is on the 'Realizar validación cruzada...' checkbox.

Figura 7.29. Interfaz para la ejecución de los algoritmos de *CKO* y *CKOE*.

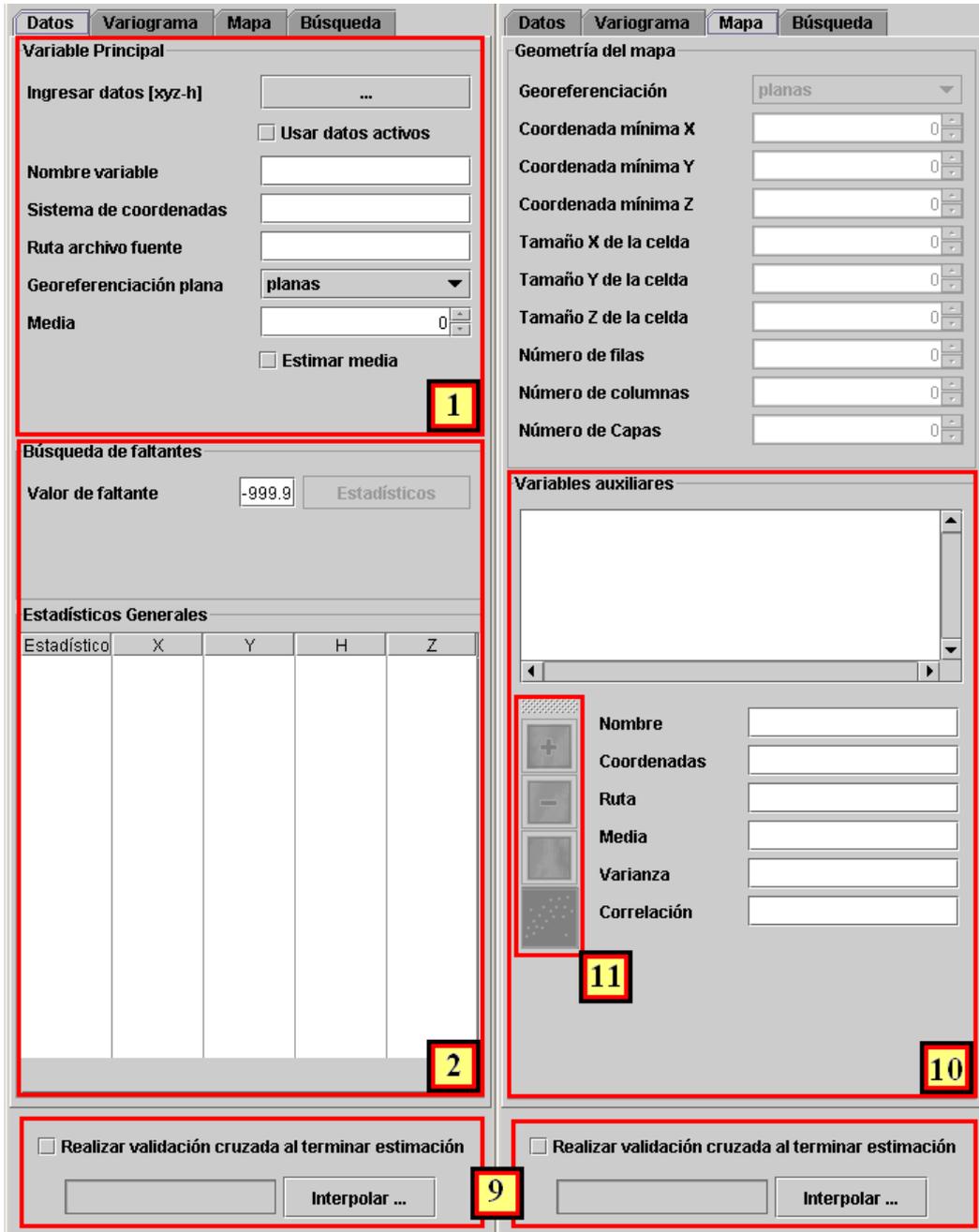


Figura 7.30. Interfaz para la ejecución del algoritmos de CKCM.

8

Ejemplo de aplicación

8.1. Introducción

En este capítulo se muestra la aplicación de ESTELLUS a un caso real de estudio. Se trata de la modelización espacial y la estimación de campos de permeabilidad utilizando la base de datos Berea (Giordano et al., 1985).

8.2. La base de datos Berea

El conjunto de datos Berea está integrado por 1600 medidas de permeabilidad distribuidas regularmente en una malla de 40 por 40 celdas sobre una sección de arenisca de Berea de 40 cm por 40 cm. Este conjunto de datos ha sido aumentado con un conjunto de 1600 medidas de porosidad generadas sintéticamente, asegurando la existencia de cierta correlación entre ambas variables. Estos campos pueden verse en la figura 8.1.

De los 1600 datos de permeabilidad se han seleccionado 64 datos por muestreo aleatorio estratificado. Tras dividir el área de trabajo en 64 paneles de igual tamaño, se ha elegido una muestra al azar en cada uno de los paneles. En cada uno de los 64 puntos se ha muestreado permeabilidad y porosidad. La representación gráfica de estos datos puede verse en la figura 8.2.

El conjunto de 1600 datos de porosidad pretende asemejarse a lo que podría obtenerse de una campaña de geofísica (sísmica, tomografía, radar). Se considera que se dispone de información acerca de una variable externa, correlacionada con la permeabilidad, en todas las localizaciones del dominio de trabajo.

Por lo tanto, la información disponible para realizar este ejercicio se compone de:

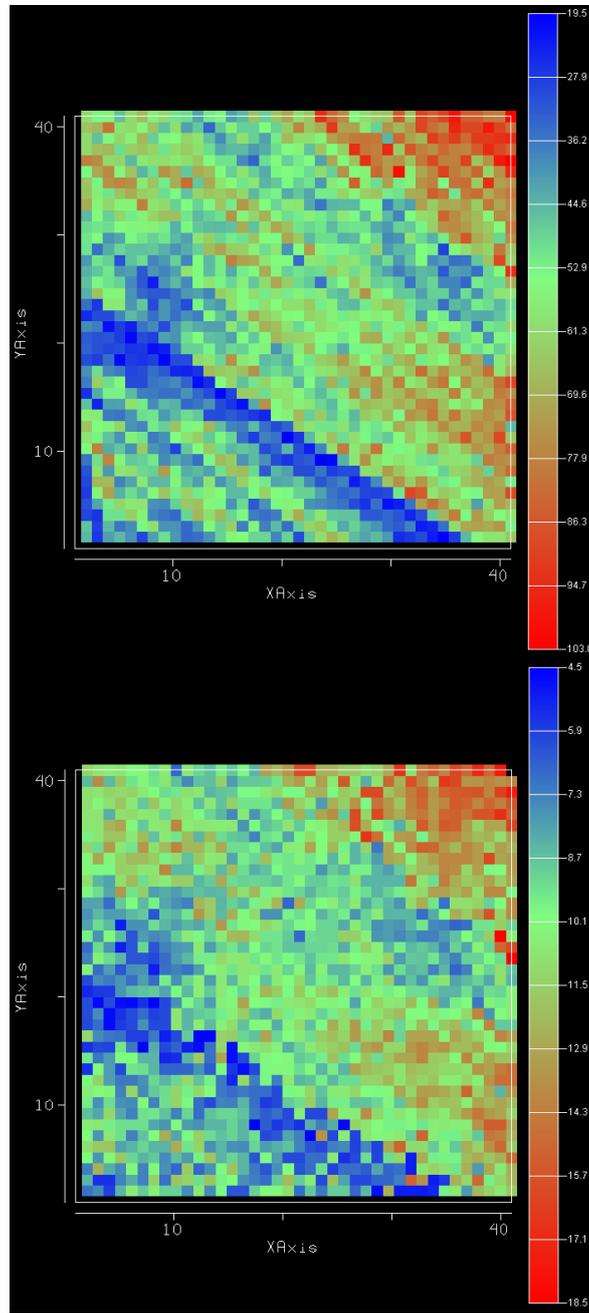


Figura 8.1. Campos de la permeabilidad (arriba) y de la porosidad (abajo) de la base de datos Berea.

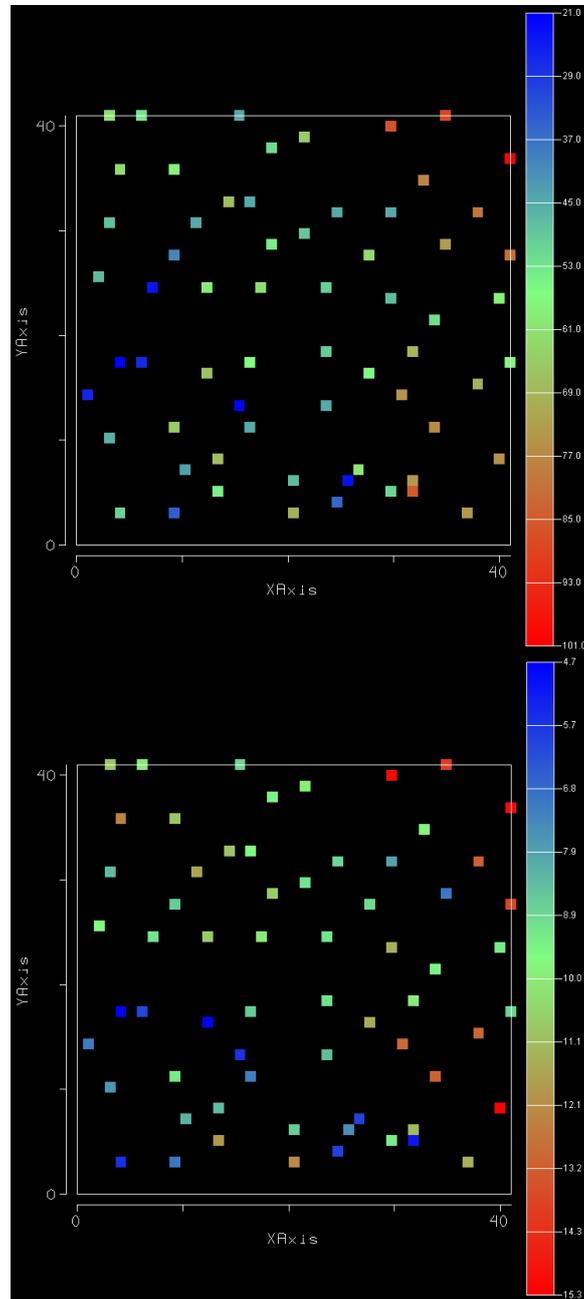


Figura 8.2. Muestreo aleatorio estratificado de la permeabilidad (arriba) y valores de la porosidad correspondientes (abajo).

- 64 medidas de permeabilidad (medidas en darcios) y
- 1600 medidas de porosidad (medidas en porcentaje).

y el objetivo del estudio es realizar el análisis descriptivo de esta información, proponer un modelo de continuidad espacial y aplicarlo para la estimación de campos de la permeabilidad vía las técnicas de interpolación implementadas en ESTELLUS.

8.3. Análisis descriptivo de la información

Puede verse en la figura 8.2 (arriba), que los mayores valores de la permeabilidad (representados por los colores más cercanos al rojo) se encuentran ubicados en la región nororiental de la zona de estudio, mientras que las zonas donde se manifiesta más baja permeabilidad se ubican en las regiones suroccidentales y occidentales (representados por los colores más cercanos a los azules oscuros). En la región central de la zona de estudio se manifiesta un comportamiento que tiende a valores medios de la permeabilidad. Para la región de estudio completa estos se ubican entre los 21 y 101. Un comportamiento análogo al anterior se visualiza para la porosidad, en la figura 8.2 (abajo), con la salvedad de que algunos valores máximos se presentan en la región suroriental de la zona de estudio. Los valores para la porosidad se encuentran entre 4.69 y 15.32. Para ambos conjuntos de datos se nota que la dirección de menor variación en las observaciones disponibles corresponde a la suroriental-noroccidental (aproximadamente a -50° de acimut, dirección $N50W$), mientras que en la dirección perpendicular se manifiesta una variación mucho más marcada de los registros disponibles (aproximadamente un acimut de $N40E$). Lo anterior se manifiesta en la variación de los colores que representan los valores de ambas variables, presentadas en la barra de colores, a la derecha. En la dirección de menor variación hay poco cambio de color, el cual se mantiene entre un amarillo claro y azul claro, mientras que en la dirección perpendicular se manifiesta variación desde los azules oscuros hasta los rojos fuertes, lo cual resulta ser análogo al cambio de las magnitudes de las observaciones de las variables analizadas, permeabilidad y porosidad. Más adelante confirmaremos estas observaciones mediante el análisis estructural.

En la figura 8.3 se presentan los histogramas, histogramas acumulados y funciones de probabilidad acumuladas calculados para la permeabilidad y la porosidad a partir de los 64 datos disponibles en la región. Estos gráficos se han generado ejecutando del módulo de descripción univariada de ESTELLUS. Se observa que para la permeabilidad se tiene gran cantidad de valores pequeños que se encuentran muy alejados del valor esperado (56.2), lo cual incrementa la varianza de la muestra. Además, gran cantidad de observaciones se encuentran entre los 45 y los 67, intervalo correspondiente al rango intercuartil. Ahora bien, para la porosidad se observa que la distribución encontrada es bastante simétrica, más aún si se considera que la media y la mediana son muy parecidas ($media = 9.6$, $mediana = 9.3$). La mayor cantidad de valores de porosidad, correspondientes a la zona del rango intercuartil se localizan entre los valores 6.16 y 11.14.

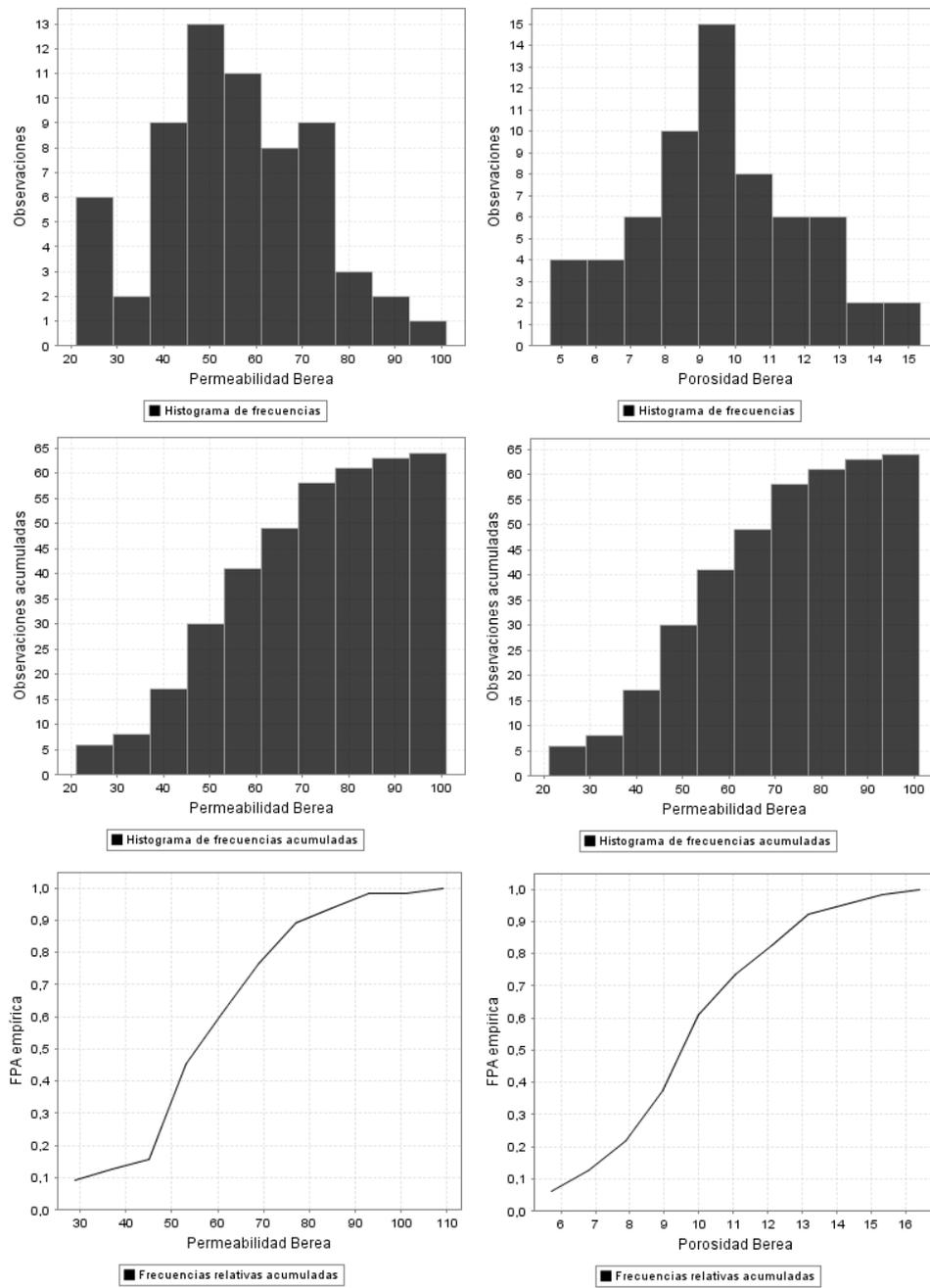


Figura 8.3. Resultados del análisis descriptivo univariado para las muestras de 64 datos de permeabilidad y porosidad.

Estadístico	Permeabilidad	Porosidad
Máximo	101.0	15.3
Mínimo	21.0	4.7
Mediana	55.0	9.3
Media geométrica	53.3	9.3
Media armónica	50.0	9.3
Media cuadrática	58.74	9.9
Media	56.2	9.6
Rango	80	10.6
Rango intercuartil	22.8	2.92
Rango percentilar	14.8	7.1
Varianza	295.3	6.4
Desviación estándar	17.2	2.5
Desviación media	13.6	2.0
Coefficiente de asimetría	0.045	0.220
Coefficiente de variación	0.31	0.26
Coefficiente de kurtosis	2.8	2.7

Tabla 8.1. Estadísticos descriptivos univariados de la permeabilidad y la porosidad para las 64 localizaciones seleccionadas en el dominio de análisis

Cuantil	Probabilidad	Permeabilidad	Porosidad
Primer decil	0.1	33.8	6.1
Segundo decil	0.2	45.0	7.5
Tercer decil	0.3	47.0	8.5
Cuarto decil	0.4	10.6	9.0
Quinto decil	0.5	55.5	9.3
Sexto decil	0.6	59.9	9.9
Séptimo decil	0.7	65.5	10.9
Octavo decil	0.8	72.1	11.6
Noveno decil	0.9	77.3	13.1
Primer cuartil	0.25	45.0	8.3
Segundo cuartil	0.5	55.5	9.3
Tercer cuartil	0.75	68.7	11.2

Tabla 8.2. Tabla de cuantiles estimados para las muestras de permeabilidad y porosidad

Otros estadísticos descriptivos de las muestras se presentan en la tabla 8.1, mientras que en la tabla 8.2 se presentan los valores de los cuantiles, que caracterizan la función

de probabilidad acumulada experimental. Aparte de las características previamente mencionadas, se encuentra que las funciones de distribución de ambas variables son bastantes simétricas lo cual se observa tanto por inspección visual de la figura 8.3 y se corrobora con el bajo valor del coeficiente de asimetría, consignado en la tabla 8.1, que ha resultado ser cercano a cero. Además, dado que el coeficiente de kurtosis es bajo, la función de probabilidad para la permeabilidad es poco puntiaguda. Por el contrario, el valor de la kurtosis obtenida para la porosidad es un poco más elevada (0.202), lo cual indica que su función de probabilidad experimental es más puntiaguda. También se observa que, puesto que el coeficiente de variación de la muestra de permeabilidad es mayor que el correspondiente para la muestra de porosidad, puede decirse que existe mayor variabilidad en la muestra de permeabilidad.

Para cuantificar las relaciones existentes entre las variables permeabilidad y porosidad se ha llevado a cabo un análisis bivariado de la información. En la figura 8.4 se presenta el diagrama de dispersión para las 64 parejas de datos de permeabilidad y porosidad. Se obtuvo que la permeabilidad y la porosidad observadas manifiestan algún grado de correlación, especialmente cuando se considera que la relación parece ser lineal ($\rho = 0.635$). Aún así, se observan algunas parejas de datos bastante alejadas de la tendencia general.

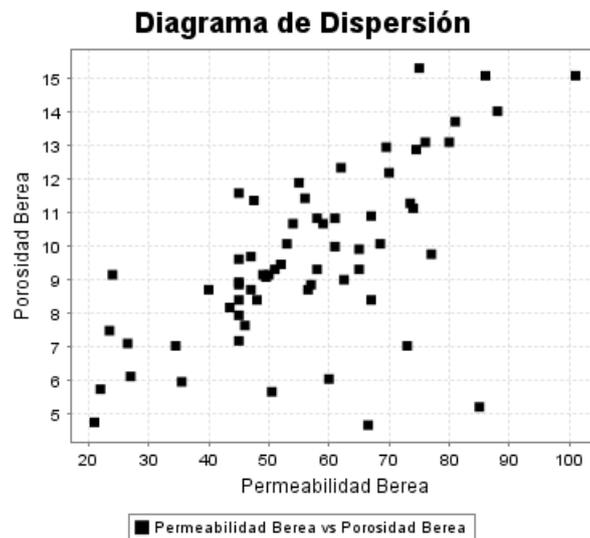


Figura 8.4. Diagrama de dispersión para las 64 parejas de datos de permeabilidad y porosidad de Berea.

8.4. Análisis estructural

El análisis estructural tiene por fin proponer un modelo matemático que explique el comportamiento espacial de los atributos registrados para estudiar un determinado fenómeno. De especial interés es la identificación de la existencia de anisotropía, lo cual se puede visualizar a través del cambio que los parámetros medidos tienen como una función de la dirección. Para encontrar las direcciones de máxima y mínima continuidad de la permeabilidad y la porosidad, se calculan los variogramas experimentales directos y el variograma cruzado entre ambas para diferentes direcciones. En la figura 8.5 se muestran tres gráficos que contienen los variogramas en las direcciones exploradas inicialmente, organizados de la siguiente manera: permeabilidad (arriba), porosidad (abajo) y cruzado entre ambas (centro). Cada uno de ellos se presenta como variograma 1,..., variograma 6 y corresponden a las direcciones N10E, N40E, N70E, N100E, N130E y N160E, respectivamente. A partir de una observación de estos gráficos, se puede establecer un intervalo de direcciones para el cual la variación es mayor y menor en el conjunto de observaciones de permeabilidad y porosidad. En este caso, dichas direcciones corresponden a aquellas para las cuales las mesetas o el valor del variograma para una distancia más o menos considerable, son máximos y mínimos. A partir de los resultados obtenidos se establece que las direcciones de mínima y máxima variación se encuentran en los intervalos (N100E; N130E) y (N40E; N70E), respectivamente.

Para definir con mayor certeza las direcciones de los ejes de anisotropía se estimaron los variogramas superficiales de la permeabilidad, la porosidad y el cruzado. Los resultados, en forma de campos raster se presentan en la figura 8.6. En este estudio se establecieron como direcciones de anisotropía las dadas por los acimuts $N123E$ y $N33E$, para la mínima y máxima variación respectivamente.

Como se comentó en el capítulo 5, el variograma es una función vectorial que mide la distancia estructural entre dos puntos en función de la distancia. En general, para procesos estocásticos espaciales de memoria finita, el variograma se manifiesta como una función creciente que se estabiliza para un valor de distancia llamada alcance, para la cual la correlación, o correlación cruzada, según sea el caso, de la(s) variable(s) desaparece(n). Usualmente su comportamiento varía con la dirección considerada. La pendiente del variograma es más suave en la dirección de máxima continuidad indicando que la correlación se pierde más lentamente, y es más abrupta en la dirección de mínima continuidad donde la correlación desaparece más rápidamente. Para la región de estudio analizada se ha seleccionado un modelo de correogionalización anisotrópico geométrico, cuyas expresiones pueden generalizarse de la siguiente manera:

$$\gamma(\mathbf{h}) = C_0 + C_1 \cdot Sph \left(\sqrt{\left[\frac{h_x}{a_{1x}} \right]^2 + \left[\frac{h_y}{a_{1y}} \right]^2} \right) \quad (8.1)$$

y está integrado por las siguientes estructuras imbricadas:

- Un efecto pepita C_0 .



Figura 8.5. Variogramas direccionales de la permeabilidad (arriba), la porosidad (abajo) y cruzado (centro) para diferentes direcciones.

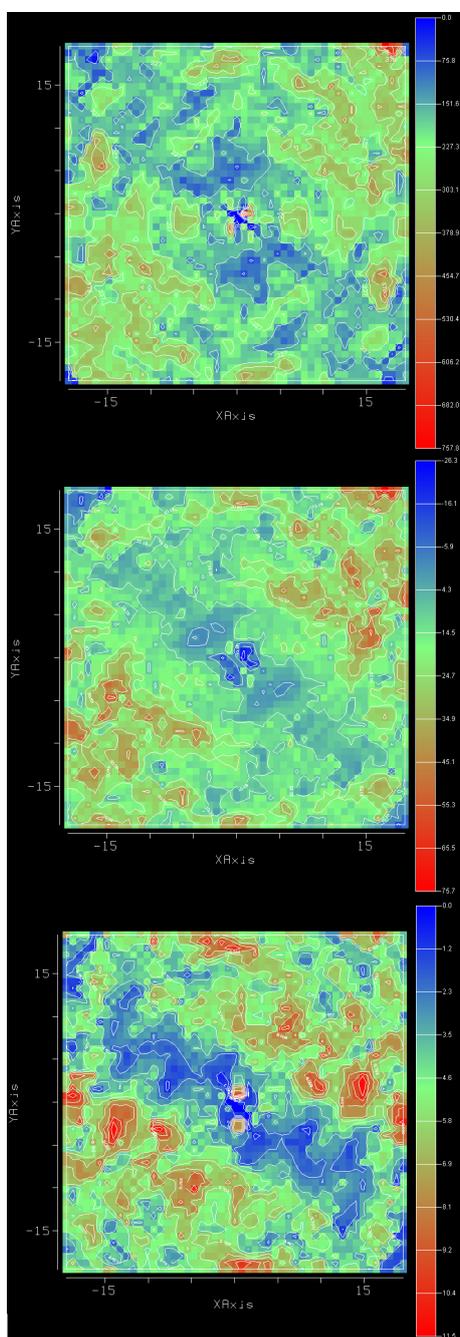


Figura 8.6. Variogramas superficiales de la permeabilidad (arriba), la porosidad (abajo) y cruzado (centro).

- Una estructura esférica anisotrópica con un rango mayor a_{1x} en la dirección h_x y un rango menor a_{1y} en la dirección perpendicular h_y y una contribución C_1 .

La dirección h_x corresponde a la dirección de máxima continuidad y la dirección h_y a la de mínima continuidad. Reemplazando los valores ajustados para los parámetros de la corregeionalización se tiene la siguiente expresión:

$$\gamma(\mathbf{h}) = \begin{bmatrix} 35 & 7.5 \\ 7.5 & 2 \end{bmatrix} + \begin{bmatrix} 240 & 18 \\ 18 & 35 \end{bmatrix} \cdot Sph \left(\sqrt{\left[\frac{h_x}{a_{1x}} \right]^2 + \left[\frac{h_y}{a_{1y}} \right]^2} \right)_{\alpha=156^\circ, a=13, b=75} \quad (8.2)$$

Los variogramas ajustados correspondientes a los parámetros presentados en la ecuación 8.2, junto con los experimentales pueden observarse en la figura 8.7, mientras que el modelo de corregeionalización, en forma de variograma superficial se observa en la figura 8.8.

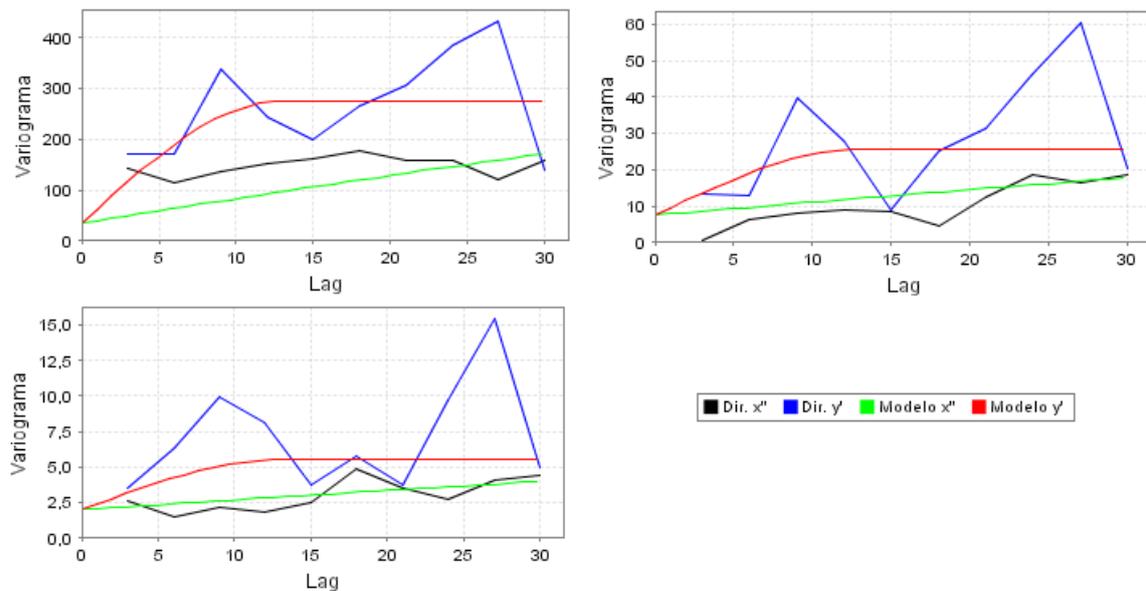


Figura 8.7. Modelo de corregeionalización para la permeabilidad (arriba izquierda), cruzado (arriba derecha) y la porosidad (abajo izquierda).

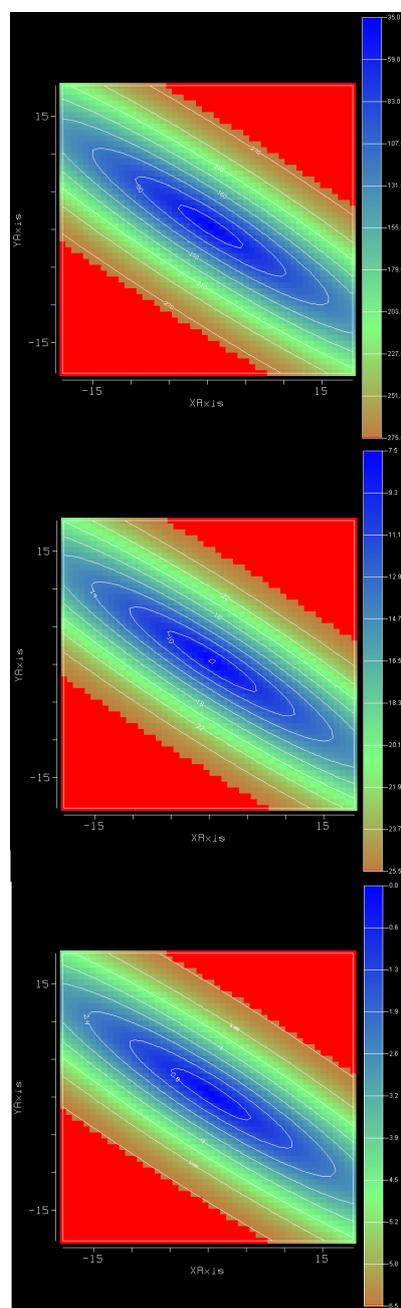


Figura 8.8. Variograma superficial de la permeabilidad (arriba), la porosidad (abajo) y cruzado (centro).

8.5. Estimación

En esta sección se presentan los resultados obtenidos de la aplicación de los algoritmos de estimación implementados en ESTELLUS. Se evalúan los algoritmos de krigeado ordinario, krigeado simple, krigeado con deriva externa, cokrigeado ordinario y cokrigeado colocalizado bajo un modelo de Markov. Para valorar los campos interpolados se han realizado los siguientes análisis:

- Una evaluación visual del campo de permeabilidad resultante, con respecto al campo de referencia.
- Una descripción univariada de los valores estimados.
- Una descripción bivariada de los valores estimados y de referencia.

Todos los campos estimados se obtuvieron utilizando una elipse de búsqueda con semieje mayor de 80 unidades y semieje menor de 60 unidades, la cual resulta ser una zona del espacio lo suficientemente grande para contener la elipse de anisotropía, girada un ángulo de $N123E$.

8.5.1. Kriging ordinario y krigeado simple

Ambas metodologías sólo tienen en cuenta la información de permeabilidad, por lo que son las aproximaciones más sencillas. En las figuras 8.9 y 8.10 se presentan los campos obtenidos. Para el caso de krigeado simple se utilizó el promedio de los valores de permeabilidad de los 64 datos como media. Comparando ambos campos con el campo original se observa que, a pesar de que muchas características locales no han sido representadas, la tendencia de variabilidad espacial ha sido adecuadamente capturada.

8.5.2. Kriging con una deriva externa

Este método es aplicable cuando, de alguna manera, se puede afirmar que la variable secundaria define la tendencia de la variable principal. En ese orden de ideas, la información secundaria formada por el campo de porosidad puede ser un buen indicador para la tendencia de la variable principal. El *KDE* requiere conocer la variable secundaria de forma exhaustiva. El campo de permeabilidad interpolado por *KDE* se muestra en la figura 8.11. Se observa que el *KDE* logra capturar la variabilidad de la información secundaria en algunas zonas donde la correlación local es más elevada entre ambas variables, junto con la tendencia espacial con respecto a la presentada en el campo original.

8.5.3. Cokrigeado ordinario estandarizado

El método de *CKOE* trata la información que se desea incorporar al procedimiento de interpolación como covariada, es decir que se utiliza la correlación espacial de las variables

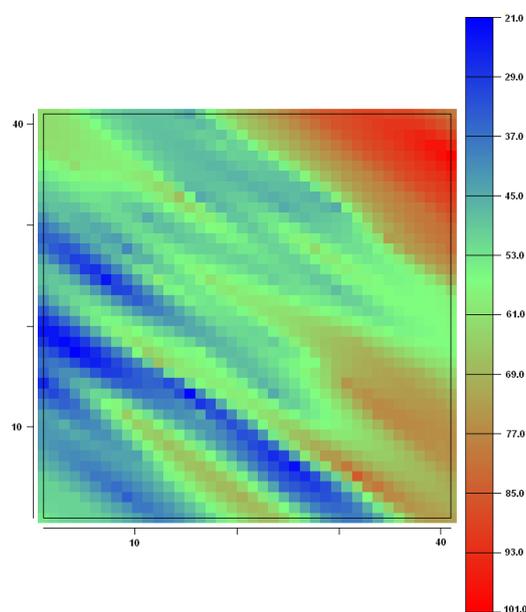


Figura 8.9. Campo de permeabilidad estimado mediante krigeado ordinario.

principal y secundarias y la correlación cruzada entre las mismas. De todas las técnicas de estimación que se usan en este trabajo, el cokrigeado es la que mayor información sobre la correlación entre las variables incorpora. Desgraciadamente, el hecho de considerar de forma amplia la correlación entre las variables integradas en la interpolación hace que deba emplearse un mayor esfuerzo de modelación de variogramas, para conseguir el modelo de corregealización. Para la aplicación de este método en la interpolación del campo de permeabilidad se utilizaron las 64 parejas de permeabilidad y porosidad disponibles. El campo de permeabilidad resultante se presenta en la figura 8.12. A simple vista, la información secundaria incluida no parece aportar mucho más a la estimación de lo que puede influir la información principal. Esto puede originarse a los pocos datos secundarios integrados en el análisis.

8.5.4. Cokrigeado colocalizado bajo un modelo de Markov

El método del *CKCM* considera como información secundaria el dato de porosidad colocalizado y propone un modelo de continuidad espacial cruzado escalando el modelo correspondiente a la permeabilidad. Lo anterior es de gran utilidad puesto que ahorra tiempo de cálculo computacional y facilita la estimación del modelo de corregealización. Análogamente al *KDE*, el *CKCM* requiere que la información secundaria se encuentre exhaustivamente muestreada. El campo de permeabilidad resultante se presenta en la figura 8.13. Por inspección visual puede verse que el campo obtenido por este método

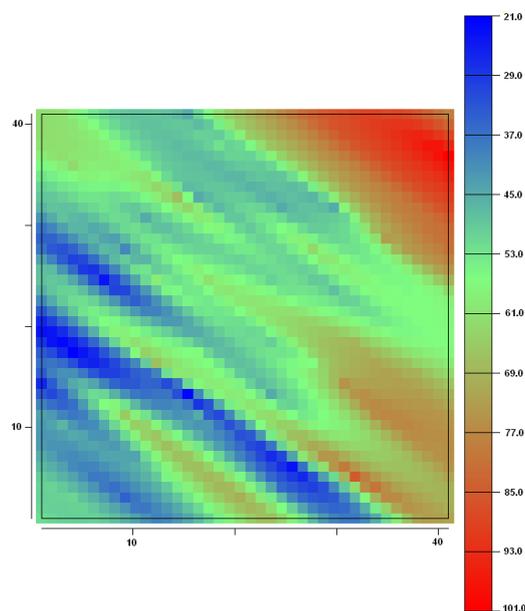


Figura 8.10. Campo de permeabilidad estimado mediante krigeado simple.

conserva bastante bien la tendencia de variabilidad espacial de la información secundaria y puede reproducir ciertos comportamiento locales que ninguno de los métodos presentados anteriormente pudo lograr.

8.6. Análisis de las estimaciones por validación cruzada

Para evaluar el comportamiento de las diferentes técnicas de estimación espacial implementadas, se ha hecho un análisis univariado de los valores estimados mediante validación cruzada. A partir de dicha validación se obtuvieron varios conjuntos de valores de permeabilidad estimados con cada técnica, los cuales se comparan con los valores de referencia, que corresponden a las 64 localizaciones con información de permeabilidad. Los estadísticos principales relacionados con las distribuciones de los valores estimados y de referencia se resumen en la tabla 8.3.

En dicha tabla se observa que las medias son muy bien reproducidas por todos los algoritmos, siendo el *KDE* el que mejor la representa. La varianza, la desviación estándar y los valores máximos son subestimados por todos los métodos de estimación. Lo anterior resulta consistente con el hecho de que, como lo comentan algunos investigadores, la subestimación de la variabilidad de las estimaciones es típica de las técnicas de interpolación que combinan varios valores para construir un estimador óptimo. Además, para el caso de las técnicas bivariadas, esto también puede atribuirse a que, como se comentó previamente,

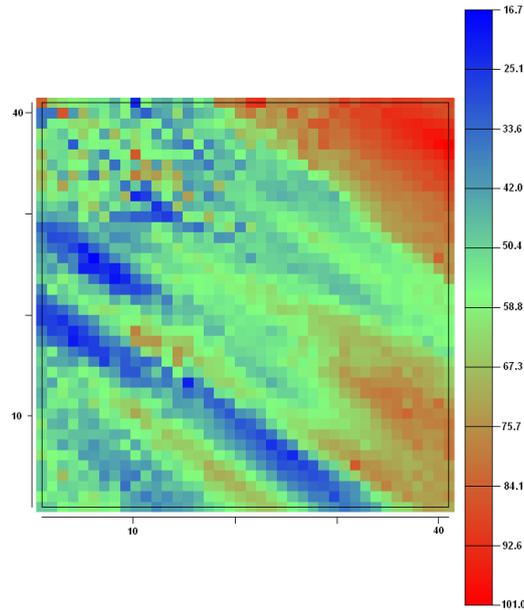


Figura 8.11. Campo de permeabilidad estimado mediante krigeado con deriva externa.

Estadístico	Datos	KO	KS	KDE	CKOE	CKCM
Max	101	85.39	94.78	95.87	94.28	96.518
Min	21	23.73	23.81	17.20	23.51	24.80
m	56.21	55.47	55.97	56.23	56.31	56.00
M	55.5	53.908	54.21	55.70	53.29	54.68
Var	295.30	180.13	201.22	216.50	279.46	226.82
σ	17.18	13.42	14.19	14.71	16.71	15.06
CV	0.31	0.24	0.25	0.26	0.3	0.27
CA	0.045	0.143	0.318	0.184	0.527	0.49
K	2.77	2.81	3.06	3.46	2.91	3.24
ρ	—	0.65	0.65	0.64	0.61	0.69

Tabla 8.3. Estadísticos derivados de la información obtenida por validación cruzada para las distintas técnicas de interpolación empleadas (Max: máximo, Min: mínimo, m: media, M: mediana, Var: varianza, σ : desviación estándar, CV: coeficiente de variación, CA: coeficiente de asimetría, K: coeficiente de kurtosis, ρ : Correlación de Pearson).

la información secundaria no logra capturar los valores extremos, máximos y mínimos. Adicionalmente, de acuerdo con el coeficiente de variación en la tabla 8.1, la variabilidad

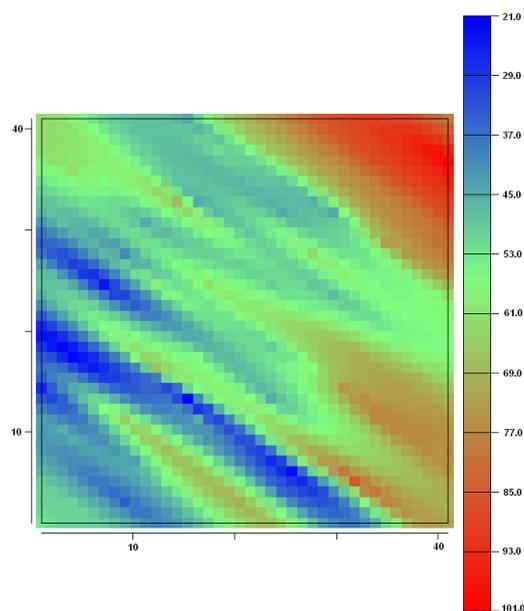


Figura 8.12. Campo de permeabilidad estimado mediante cokrigado ordinario.

de la porosidad es menor que la de la permeabilidad, al menos con los valores disponibles en las 64 localizaciones, con lo cual la tendencia espacial de apoyo será restringida.

Con respecto a los valores mínimos, puede observarse que ninguna técnica produce resultados satisfactorios, todas ellas sobreestiman dicho valor, con excepción del *KDE* que lo subestima por mucho. En general, el valor mínimo estimado está cerca de 25. El coeficiente de variación y el coeficiente de asimetría son mejor reproducidos por el algoritmo de *KDE* seguido por *CKOE*. Igualmente, en la figura 8.14 se presentan los histogramas de las muestras de permeabilidad estimadas. En estos gráficos se observa que valores más cercanos se estiman mediante las técnicas de *CKOE* y *CKCM*. En general, se observa que la zona central del histograma, para observaciones de permeabilidad cercanas a 70 *mm/día* ninguno de los algoritmos puede llegar a capturar las frecuencias estimadas de acuerdo con las reales, como se observa en el histograma superior. Dichos valores parecen ser subestimados, por lo cual las funciones de probabilidad empírica obtenidas son más puntiagudas, como revelan los valores de kurtosis.

Un diagrama de dispersión provee información adicional acerca del comportamiento de los algoritmos empleados. Bajo condiciones ideales, es decir, si los valores estimados y de referencia son iguales, se esperaría que los valores estimados y los verdaderos se ajustaran a una línea recta con pendiente de 45° si la escala de los ejes es la misma. Los diagramas de dispersión de los valores verdaderos y estimados por los diferentes métodos se representan en la figura 8.15. En ellos se observa que todas las metodologías arrojan

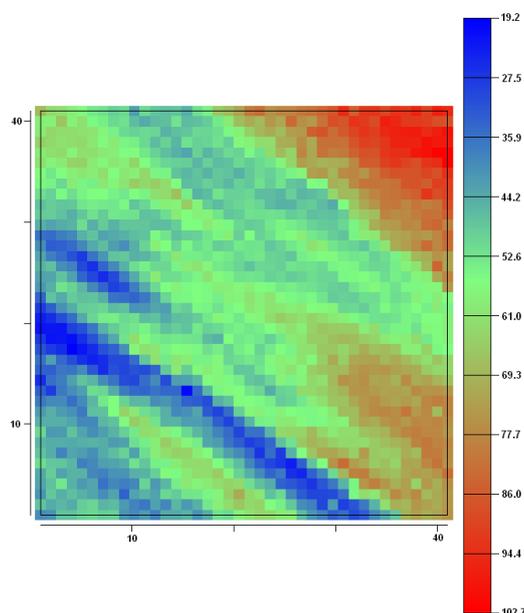


Figura 8.13. Campo de permeabilidad estimado mediante cokrigado colocalizado bajo un modelo de Markov.

diagramas de dispersión similares, con lo cual los valores del coeficiente de correlación de Pearson también lo son, como se observa en la fila final de la tabla 8.3. Podríamos decir que los valores mejor reproducidos corresponden a la técnica *CKCM*.

8.7. Comentarios finales

En este capítulo se ha presentado una aplicación completa de un estudio geoestadístico de estimación espacial utilizando las herramientas diseñadas para ESTELLUS. La herramienta construida es efectiva tanto para realizar el procedimiento de estimación como para la presentación y análisis de los resultados. El caso comentado es un ejemplo sintético, en el cual las coordenadas de las entidades puntuales con información están referidas a un sistema local de coordenadas espaciales, más no se encuentran debidamente georeferenciadas.

Aplicaciones más detalladas de casos reales empleando los algoritmos de estimación usados para ESTELLUS pueden encontrarse en el trabajo de Álvarez (2007), quien construyó un nuevo campo de precipitación para Colombia usando información principal de estaciones climatológicas e información de radar proveniente de la *TRMM* (Kummerow et al., 1998). Esta aplicación solamente utilizó los algoritmos computacionales de la bibliote-

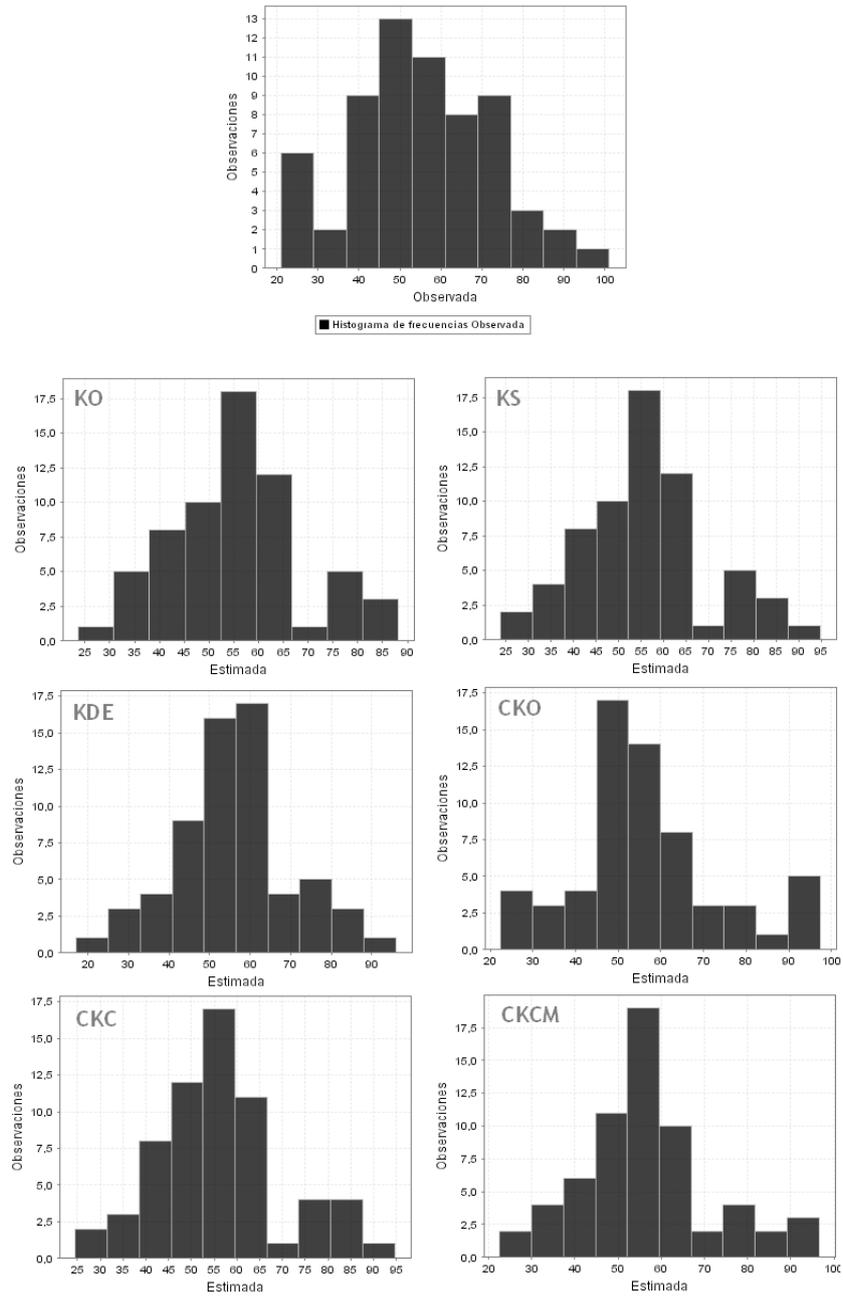


Figura 8.14. Histogramas para los valores de la permeabilidad estimados por validación cruzada.

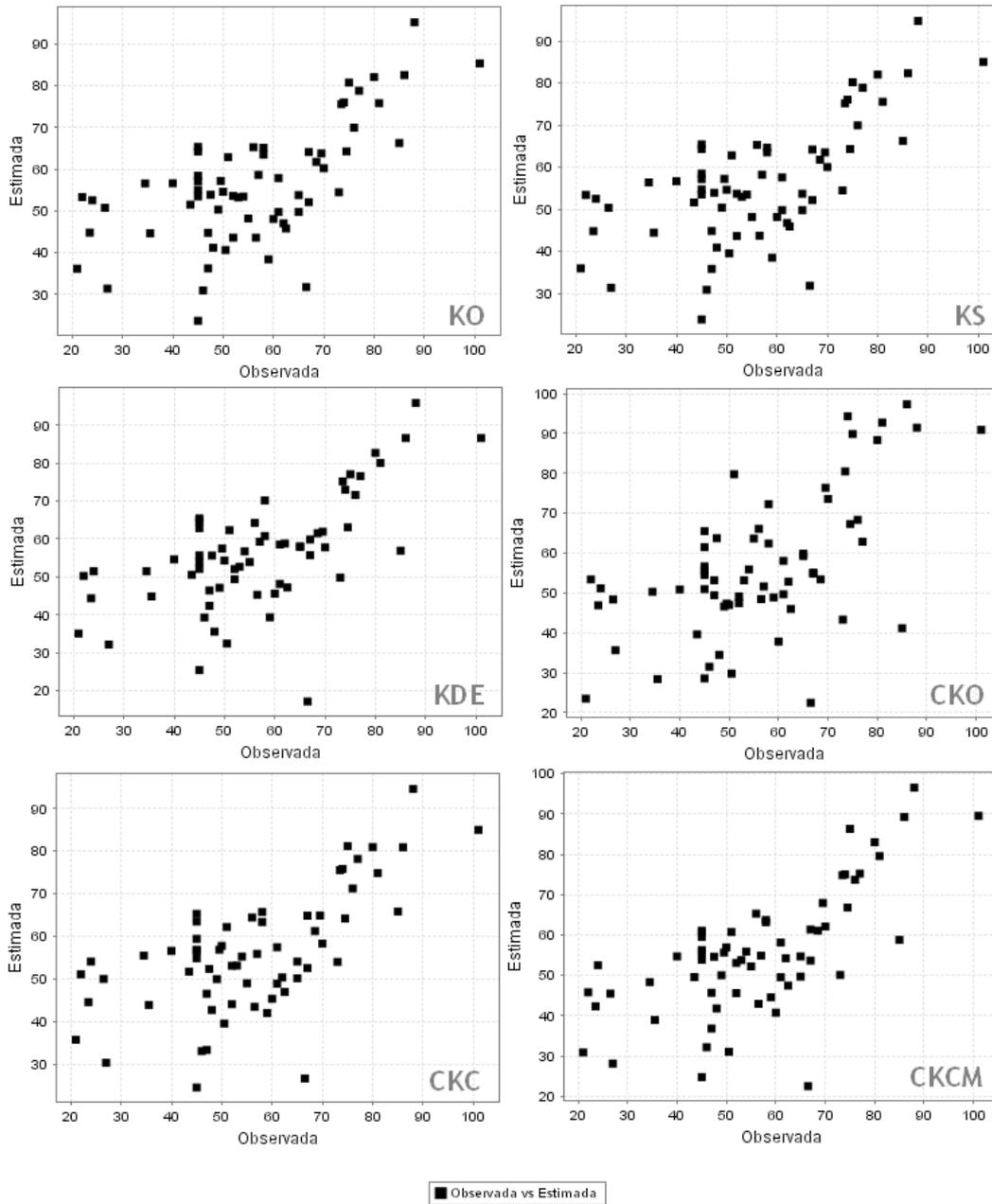


Figura 8.15. Diagramas de dispersión de los valores de la permeabilidad estimados por validación cruzada vs. los valores registrados.

ca de clases, puesto que en ese momento no se disponía de una interfaz gráfica de usuario de buena calidad.

9

Conclusiones

La Geoestadística es una ciencia que aporta herramientas que ayudan a los analistas de las ciencias de la tierra en los procedimientos de caracterización, extracción de patrones de comportamiento espacial, estimación y modelación de la incertidumbre de variables espacialmente distribuidas. Este trabajo es la implementación de una librería que contiene dichos procedimientos, junto con un interfaz gráfica de usuario, escritas ambas en lenguaje de programación `Java` y acoplada con una plataforma de visualización gratuita. Como se trata de una librería bajo los términos del software libre, quien lo desee puede adaptar el código a sus necesidades.

Durante la revisión bibliográfica se encontraron diferentes tipos de herramientas que presentaban algunas ventajas y/o desventajas unas con respecto a las otras. En el desarrollo de ESTELLUS se intentó reunir las ventajas encontradas en las bibliotecas geoes-tadísticas existentes, intentando adoptar las características más convenientes de acuerdo con el propósito general del trabajo. En este orden de ideas, a lo largo del documento se han discutido las características de ESTELLUS, de las cuales las principales son:

- El código ha sido escrito utilizando programación orientada a objetos espaciales para poder representar tanto las características y atributos de las variables, como para definir explícitamente lo relacionado con su localización espacial, sistema de coordenadas utilizado y georeferenciación asociada.
- Las librerías de programas se construyeron en lenguaje de programación `Java`, el cual es multiplataforma, por lo cual no existen problemas de instalación del software que tengan que ver con el procesador o con el sistema operativo que este utilizando el usuario. Adicionalmente, un código en `Java` es fácilmente comprendido por

programadores con nivel intermedio, por lo cual el software podrá mantenerse en crecimiento continuo.

- El código ha sido creado de forma modular y ha sido extensamente documentado, por lo cual se facilita su comprensión. Por un lado, se tiene la biblioteca de algoritmos con los cuales un programador avanzado puede desarrollar aplicaciones nuevas o efectuar gran cantidad de cálculos o estimaciones; mientras que por el otro lado se tienen las aplicaciones en forma de ventanas que permiten al usuario general efectuar interactuar con el software de forma sencilla.
- La librería de clases programada aprovecha las ventajas de diferentes bibliotecas de clases **Java** de uso libre bajo licencia *GPL*, como lo son **JAMA** para la manipulación de matrices, **HIDROSIG** para el manejo de los campos raster bidimensionales, la librería del Dr. Flanagan para el ajuste automático del variograma, **VISAD** para el despliegue gráfico de la información y resultados y **JFREECHART** para la visualización de histogramas, variogramas, diagramas de dispersión y gráficos de series. En su mayoría los códigos usados han sido adaptados de acuerdo con los requerimientos de los algoritmos de estimación implementados en **ESTELLUS**.
- El hecho de tener en cuenta el atributo espacial de las entidades geográficas permite que los resultados estimados sean compatibles con los sistemas de información geográfica. Por ello se integran algoritmos de importación y exportación desde y hacia los formatos *SIG* más frecuentemente utilizados en la comunidad científica.
- **ESTELLUS** ha sido creado como un software geoestadístico, pero con la compatibilidad suficiente para que pueda ser usado efectivamente en cualquier proyecto relacionado con modelación de sistemas naturales. Las librerías computacionales desarrolladas soportan el análisis de problemas bidimensionales y tridimensionales, con lo cual la aplicación del software puede realizarse en diferentes áreas de las ciencias de la tierra, como lo son la hidrología superficial, la climatología, la meteorología, la hidrogeología, la ciencia del petróleo, la geología, las ciencias agrícolas, las ciencias del suelo, etc.
- **ESTELLUS** posee una extensa librería para el análisis descriptivo de las muestras de información a utilizar, herramienta útil para describir tendencias y filtrar información, de forma que se obtenga un conjunto de datos apropiados y robustos que describan claramente las tendencias espaciales del atributo a analizar o representar.
- La librería de aplicaciones para el análisis estructural de **ESTELLUS** contiene gran cantidad de funciones para detectar tendencias espaciales de variabilidad, extracción automática y múltiple de diferentes tipos de variograma experimentales; junto con técnicas de ajuste automático y visual de modelos. Así, es una herramienta flexible que permite al usuario la exploración de diferentes estrategias de análisis estructural. Adicionalmente incorpora utilidades para manipulación de la información de forma

externa al software, como el almacenamiento en archivos de texto de los variogramas, con lo cual el usuario puede modificar los resultados sin tener que usar la interfaz de forma rígida. Lo anterior también va en pro de la eficiencia para manipular los resultados.

- ESTELLUS ofrece una amplia biblioteca de algoritmos de estimación espacial estocástica y determinista, como lo son vecino cercano, inverso de la distancia a un exponente, krigeado ordinario, krigeado simple, krigeado con deriva externa, cokrigeado ordinario, cokrigeado ordinario estandarizado, y cokrigeado colocalizado bajo un modelo de Markov. Adicionalmente se han diseñado herramientas para el análisis y comparación cualitativa y cuantitativa de los resultados obtenidos como los son el análisis estadístico de muestras generadas mediante validación cruzada y la manipulación interactiva de los campos estimados.
- Los algoritmos de análisis y estimación están contenidos en una sencilla pero completa interfaz del usuario, lo cual permite el ingreso a las diferentes herramientas desarrolladas y el análisis interactivo de los resultados obtenidos.

9.1. Líneas futuras de trabajo e investigación

ESTELLUS es el primer paso hacia el diseño y construcción de una aplicación informática capaz de ser utilizada para un estudio geoestadístico completo que incorpore no sólo algoritmos clásicos sino también las técnicas más novedosas. Para ello se han fijado las siguientes tareas futuras:

- Implementación de técnicas de evaluación de la incertidumbre local mediante el uso de las aproximaciones multi-gaussianas e indicadoras.
- Implementación de técnicas de evaluación de la incertidumbre espacial o simulación estocástica.
- Implementación de algoritmos para el ajuste automático de modelos de correogionalización lineal.
- Mejora de los algoritmos para el cálculo de variogramas superficiales.
- Implementación del cálculo automático de funciones no paramétricas de continuidad espacial.
- Implementación de herramientas de postprocesado de imágenes derivadas de los algoritmos de estimación y simulación.
- Implementación de algoritmos de geoestadística multipuntual.
- Incorporación de herramientas para el tratamiento de fenómenos dependientes del tiempo.

- Implementación de herramientas vectoriales que pueden facilitar la adquisición de información, como por ejemplo la importación de *shapefiles* puntuales de ESRI.

Simultáneamente, ESTELLUS requiere de una verificación, depuración y puesta a punto permanente para contrastar su robustez y amigabilidad, lo que se conseguirá aplicándolo a muchos casos de estudio. Conseguida una estabilidad moderada ESTELLUS será “presentado en sociedad” con el fin de convertirse en una herramienta útil para la comunidad científica.

Referencias bibliográficas

- Almeida, A.S. y A.G. Journel (1994) Joint simulation of multiple variables with a Markov type corregeonization model. *Mathematical Geology* 26(5), 565-588.
- Álvarez O.D. (2007) Cuantificación de la Incertidumbre en la Estimación de Campos Hidrológicos. Aplicación al Balance Hídrico de Largo Plazo. Tesis de Maestría. Universidad Nacional de Colombia.
- Álvarez, O.D. y J.E. Alarcón (2004) Evaluación de la Amenaza al Deslizamiento Superficial Empleando Sistemas de Información Geográfica. Trabajo Dirigido de Grado. Universidad Nacional de Colombia.
- Armstrong, M. y R. Jabin (1981) Variogramas models must be positive-definite. *Mathematical Geology* 13(5), 455-459.
- Barco, O.J. y L.A. Cuartas (1998) Estimación de la Evaporación en Colombia. Trabajo Dirigido de Grado. Universidad Nacional de Colombia.
- Bosque-Sendra, J. (1997) *Sistemas de Información Geográfica*. Segunda Edición. Ediciones RIALP. Madrid.
- Budd, T. (1994) *Introducción a la Programación Orientada a Objetos*. Addison-Wesley Iberoamericana. Reading.
- Burrough, P.A. (2001) GIS and geostatistics: essential partners for spatial analysis. *Environmental and Ecological Statistics* 8, 361-377.
- Burrough, P.A. y R.A. McDonnell (1998) *Principles of Geographical Information Systems*. Oxford University Press. Spatial Information and Geostatistics Series. New York.
- Carr, J.R. y K. Mela (1998) Visual Basic programs for one, two or three-dimensional geostatistical analysis. *Computers and Geosciences* 24(6), 531-536.
- Cassiraga, E.F. (1999) Incorporación de Información Blanda para la Cuantificación de la Incertidumbre: Aplicación a la Hidrogeología. Tesis Doctoral. Universidad Politécnica de Valencia.
- Chilès, J.P. y Delfiner, P. (1999) *Geostatistics. Modeling Spacial Uncertainty*. Wiley Interscience Publication. New York.

- Christakos, G. (1984) On the problem of permissible covariance and variogram models. *Water Resources Reserach* 20(2), 251-265.
- Chu, J. (1993) Contional Fractal Simulation, Sequential Indicator Simulation and Interactive Variogram Modeling. Tesis Doctoral. Stanford University.
- Chua, S.H. y R.L. Bras (1982) Optimal estimators of mean areal precipitation in regions of orographic influence. *Journal of Hydrology* 57, 23-48.
- Cressie, N. (1985) Fitting varogram models by weighted less squares. *Mathematical Geology* 17(5), 563-586.
- Declercq, F.A.N. (1986) Interpolation methods for scattered sample data: accuracy, spatial patterns, processing time. *Cartography and Geographic Information Systems* 23(3), 128-144.
- Deutsch, C.V. y A.G. Journel (1998) *GSLib, Geostatistical Software Library and User´s Guide*. Oxford University Press. New York, NY.
- Dubrulle, O. (1983) Two methods with different objective: splines and kriging. *Mathematical Geology* 15(2), 245-257.
- EPA, United States Environmental Protection Agency (1991) GEO-EAS 1.2.1 User´s Guide. En Internet: www.epa.gov/ada/download/models/geoeas.pdf
- Eriksson, M., y P. Siska (2000). Understanding anisotropy calculations. *Mathematical Geology* 32(6), 683-699.
- ESRI (2001) ARCGISTM Geostatistical Analyst: Statistics Tools for Data Exploration, Modeling and Advanced Surface Generation. ESRI©White Paper. En Internet: www.ESRI.com
- Faulkner, B.R. (2002) Java clases for nonprocedural variogram fitting. *Computers and Geosciencias* 28(2002), 387-397.
- Fernández, B. (2002) *Geodesia Para Ingenieros*. Segunda Edición. Editorial Universidad Francisco José de Caldas. Bogotá.
- Fernández-Coppel, I.A. (sin fecha) Las Coordenadas Geográficas y la Proyección UTM. Departamento de Ingeniería Agrícola y Forestal. Universidad de Valladolid.
- Flanagan, M.T. (2008) Michael Thomas Flanagan´s Java Scientific Library. En internet: <http://www.ee.ucl.ac.uk/~mflanaga/java/>.
- García-Merayo, F. (1999) *Lenguaje de Programación Fortran 90*. Editorial Paraninfo. Madrid.

- Getis, A. (1999) Spatial statistics. En: *Geographical Information Systems: Principles and Technical Issues* Volume 1. Longley, P.A., M.F. Goodchild, D.J. Maguire, y D.W. Rhind Eds. Capítulo 1, 239-251.
- Giordano, R. M., Salter, S. J. y Mohanty, K. K. (1985) The effect of permeability variations on flow in porous media. *SPE* 14365.
- Gómez-Hernández, J. (1996) Geostatistics and hydrology: an overview. *Second Joint Conference and Exhibition on Geographical Information* Volumen 1, 457-466. IOS Press. Amsterdam.
- Gómez-Hernández, J., E. Cassiraga, C. Guardiola-Albert y J. Álvarez (2000) Incorporating information from a digital elevation model for improving areal estimation of rainfall. *GeoENV III, Geostatistic For Environmental Applications*. Editado por Monestiez et al. (2001). Kluwer Academic Publishers. Berlin.
- Goovaerts, P. (1997) *Geostatistics for Natural Resources Evaluation*. Oxford University Press. Applied Geostatistics Series. New York.
- Goovaerts, P. (1999) Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology* 118, 113-129.
- Hibbard, W.L. (1995) Visualizing Scientific Computation: A System Based on Latticed Structured Data and Display Models. Doctoral dissertation. University of Wisconsin.
- Hibbard, W.L. (1998) VisAD: connecting people to computations and people to people. *Computer Graphics* 32(3), 10-12.
- Hibbard, W.L. (2002) Building 3-D user interface components using a visualization library. *Computer Graphics* 36(1), 4-7.
- Instituto Geográfico Agustín Codazzi IGAC (2004). Tipos de Coordenadas Manejados en Colombia. Publicación IGAC.
- Isaaks, E.H. y R.M. Srivastava (1989). *Applied Geostatistics*. Oxford University Press. Applied Geostatistics Series. New York.
- JFree (2005) JFreeChart: A Java Chart Library. En Internet: www.jfree.org/jfreechart/.
- Journel, A.G. (1983) Nonparametric estimation of spatial distributions. *Mathematical Geology* 15(3), 445-468.
- Journel, A.G. y C. Huijbregts (1978) *Mining Geostatistics*. Reimpresión de la quinta edición (1991) publicada en 2003. The Blackburn Press. New Jersey.
- Kennedy, M. (2000) Understanding Map Projections. ESRI publication. En Internet: www.kartografie.nl

- Kummerow, C., W. Barnes, K. Toshiaki, J. Shiue, J. y J. Simpson (1998) The tropical rainfall measuring mission (TRMM) Sensor Package. *Journal Of Atmospheric and Oceanic Technology* 15, 809-817.
- Lam, N.S. (1983) Spatial interpolation methods: A review. *The American Cartographer* 10(2). 129-149.
- Leavesley, G.H. (1997) The Modular Modeling System (MMS): A modeling framework for multidisciplinary research and operational applications. Cap. 8 de Freeman, G.E. y A.G. Frazier Eds., Proceedings of the Scientific Assessment and Strategy Team Workshop on Hydrology, Ecology, and Hydraulics, v. 5 of Kelmelis, J.A., ed., *Science for Floodplain Management into the 21st Century*, 27-31. Government Printing Office. Washington, D.C.
- Luenberger, D.G. (1969) *Optimization by Vector Space Methods*. Jhon Wiley and Sons, New York, 1969
- Mantilla, R. y V. Gupta (2005) A GIS framework to investigate the process basis for scaling statistics on river networks. *IEEE Geoscience and Remote Sensing Letters* 2(4), 404-408.
- Markstrom, S.L., G. McCabe y O. David (2002) Web-distribution of geoscientific models. *Computers and Geosciences* 28(2002), 577-581.
- Martín-Asín, F. (1990) *Geodesia y Cartografía Matemática*. Tercera Edición. Editorial Paraninfo. Madrid.
- Matheron, G. (1965) *Les Variables Regionalices et Leur Estimation*. Masson. Paris.
- Matheron, G. (1971) The Theory of Regionalized Variables and Its Applications. Cahiers du Centre de Morphologie Mathématique, Ecole des Mines, Fontainebleau.
- McDonald, M.G. y A.W. Harbaugh (1984) A Modular Threedimensional Finite-Diference Groundwater Flow Model. US Geological Survey Open-File Report, 83-875.
- Meertens, C., L.D. Murray y J. McWhirter (2004) Collaborative visualization and analysis of multi-dimensional, time-dependent and distributed data in the geosciences using the Unidata Integrated Data Viewer. EOS Transactions American Geophysical Union 85(47), Fall Meeting Supplement.
- Meertens, C., L. Estey, S. Wier, G. Bensen, A. Wahr, A. Murray y J. McWhirter (2005) Visualization of earth science data. EOS Transactions American Geophysical Union 86(52), Fall Meeting Supplement.
- Meertens, C. y S. Wier (2006) Visualization using the GEON Integrated Data Viewer (IDV). Earthscope/USArray Imaging Science and CIG Seismology Workshop. St. Louis.

- Meijernik, A., H. de Bouer, C. Mannaerts y C. Valenzuela. (1994) *Introduction to the Use of Geographical Information Systems for Practical Hydrology*. ITC Publication No 23.
- Montgomery, D.C. y G.C. Runger (1996) *Probabilidad y Estadística Aplicadas a la Ingeniería*. McGraw Hill. Buenos Aires.
- Müller, W.G. (1999) Least-squares fitting from the variogram cloud. *Statistics and Probability Letters* 43(1), 93-98.
- Nelder, J.A. y R. Mead (1965) A simplex method for function minimization. *Computer Journal* 7, 308-313.
- Olea, R. (1991) *Geostatistical Glossary and Multilingual Dictionary*. Oxford University Press. New York.
- Pardo-Igúsquiza, E. (1999) VARFIT: a fortran-77 program for fitting variogram models by weighted least squares. *Computer and Geosciences* 25(1999), 251-261.
- Pebesma, E.J y C.G. Wesseling (1998) GSTAT: a program for geostatistical modeling, prediction and simulation. *Computer and Geosciences* 24(1), 17-31.
- Pebesma, E.J. (2004) Multivariate geostatistics in S: the GSTAT package. *Computer and Geosciences* 30(2004), 683-691.
- Poveda, G., O.J. Mesa, J.I. Vélez, R. Mantilla, J.M. Ramírez, O.O. Hernández, A.F. Borja y J.A. Urzola. (2007) HidroSIG: an interactive digital atlas of Colombia's hydroclimatology. *Journal of Hydroinformatics* 9(2), 145-156.
- Poveda, G., J.I. Vélez, O.J. Mesa, A. Cuartas, J. Barco, R.I. Mantilla, J.F. Mejía, C.D. Hoyos, J.M. Ramírez, L.I. Ceballos, M.D. Zuluaga, P.A. Arias, B.A. Botero, M.I. Montoya, J.D. Giraldo y D.I. Quevedo (2007) Linking long-term water balances and statistical scaling to estimate river flows along the drainage network of Colombia. *Journal of Hydrologic Engineering* 12(1), 145-156.
- Press, W.H., S.A. Teukolsky, W.A. Vetterling y B.P. Flannery (1988) *Numerical Recipes in C*. Segunda Edición. Cambridge University Press.
- Ramírez, J.M. (2002) Extracción Automática de Redes de Drenaje a Partir de Modelos Digitales de Terreno. Trabajo Dirigido de Grado. Universidad Nacional de Colombia.
- Remy, N. (2001) GsTL: The Geostatistical Template Library in C++. Stanford Center For Reservoir Forecasting. Department of Petroleum Engineering. Stanford University.
- Remy, N. (2004) Geostatistical Earth Modeling Software: User's Manual. Stanford Center For Reservoir Forecasting. Department of Petroleum Engineering. Stanford University.

- Rühaak, W. (2006) A Java application for quality weighted 3D interpolation. *Computer and Geosciences* 32(2006), 43-51.
- Schildt, H. (2001) *Java 2. Manual de Referencia*. Cuarta Edición. McGrawHill. Madrid.
- Snyder, J.P. (1987) *Map Projections: A Working Manual*. U.S. Geological Survey Professional Paper 1395. Washington.
- Spigel, M.R. y L.J. Stephens. (2002) *Estadística*. Tercera Edición. McGraw Hill. Madrid
- UPME (2000). Atlas Hidrológico de Colombia. Informe Final. Postgrado en Aprovechamiento de Recursos Hidráulicos. Universidad Nacional de Colombia.
- Vélez, J.I., G. Poveda y O.J. Mesa (2000). Balances Hidrológicos de Colombia. Postgrado en Aprovechamiento de Recursos Hidráulicos. Universidad Nacional de Colombia.
- Wier, S. (2008) The GEON Integrated Data Viewer. CSIG 2008: Cyberinfrastructure Summer Institute for Geoscientists, San Diego.
- Wilkinson J.H y C. Reisch (1972) *Handbook for Automatic Computation, Volume 2: Linear Algebra*. Springer-Verlag, Berlin.
- Wingle, W.L, E.P. Poeter, E.P. y McKeena, S.A. (1999) UNCERT: geostatistics, uncertainty analysis and visualization software applied to groundwater flow and contaminant transport modeling. *Computer and Geosciences* 25(1999), 365-376.
- Xu, W., T. Tran, R. M. Srivastava y A. G. Journel (1992) Integrating Seismic Data in Reservoir Modelling: the Colocated Cokriging Alternative. SPE paper No 24742.
- Yin, Z. (2001) Design and Implementation of Some GsTL Foundation Classes. Stanford Center for Reservoir Forecasting. Department of Petroleum Engineering. Stanford University.
- Yin, Z. (2002) A General Kriging Framework: Related Classes and Programs in GsTL. Stanford Center For Reservoir Forecasting. Department of Petroleum Engineering. Stanford University.
- Zhang, X.F., J.C Van Eikeren y A.W. Heemink (1995) On the weighted least-squares method for fitting a semivariogram model. *Computer and Geosciences* 21(4), 605-608.
- Zheng, C. (1990) *A Modular Three-dimensional Transport Model*. United States Environmental Protection Agency and S.S. Papadopoulos and Associates, Rockville.