

# Reconocimiento de Patrones en Simulación Geoestadística

Oscar Francisco Peredo Andrade

Presentación para optar al Título de Ingeniero Civil en Computación  
Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas

21 de Noviembre de 2008

# Esquema

- 1 **Introducción**
- 2 **Antecedentes**
- 3 **Trabajo realizado**
- 4 **Resultados**
- 5 **Conclusiones y Trabajo a futuro**

# Esquema

- 1 **Introducción**
- 2 Antecedentes
- 3 Trabajo realizado
- 4 Resultados
- 5 Conclusiones y Trabajo a futuro

# Introducción

- Proyecto Fondecyt:

***Evaluación de Yacimientos mediante Simulación Estocástica integrando Estadísticas de Múltiples Puntos***

- ¿Evaluación de Yacimientos?

Herramientas: ***Geoestadística***

- ¿Simulación Estocástica?

Herramientas: ***Kriging, Simulación Convencional y no Convencional***

- ¿Estadísticas de Múltiples Puntos?

Herramientas: ***Patrones 2D***

# Introducción

- Proyecto Fondecyt:

***Evaluación de Yacimientos mediante Simulación Estocástica integrando Estadísticas de Múltiples Puntos***

- ¿Evaluación de Yacimientos?  
Herramientas: **Geoestadística**

- ¿Simulación Estocástica?  
Herramientas: *Kriging, Simulación Convencional y no Convencional*

- ¿Estadísticas de Múltiples Puntos?  
Herramientas: *Patrones 2D*

- Proyecto Fondecyt:

***Evaluación de Yacimientos mediante Simulación Estocástica integrando Estadísticas de Múltiples Puntos***

- ¿Evaluación de Yacimientos?

Herramientas: ***Geoestadística***

- ¿Simulación Estocástica?

Herramientas: ***Kriging, Simulación Convencional y no Convencional***

- ¿Estadísticas de Múltiples Puntos?

Herramientas: ***Patrones 2D***

- Proyecto Fondecyt:

***Evaluación de Yacimientos mediante Simulación Estocástica integrando Estadísticas de Múltiples Puntos***

- ¿Evaluación de Yacimientos?

Herramientas: ***Geoestadística***

- ¿Simulación Estocástica?

Herramientas: ***Kriging, Simulación Convencional y no Convencional***

- ¿Estadísticas de Múltiples Puntos?

Herramientas: ***Patrones 2D***

# Esquema

- 1 Introducción
- 2 Antecedentes**
- 3 Trabajo realizado
- 4 Resultados
- 5 Conclusiones y Trabajo a futuro



- Rama de la Estadística que pone énfasis en contexto geológico y espacial de los datos.
- Generación de modelos de bloques de leyes 3D para planificación minera.
- Estimación de las reservas locales y globales.
- Cuantificación de incertidumbre en contenido y predicción de la producción.

- Rama de la Estadística que pone énfasis en contexto geológico y espacial de los datos.
- Generación de modelos de bloques de leyes 3D para planificación minera.
- Estimación de las reservas locales y globales.
- Cuantificación de incertidumbre en contenido y predicción de la producción.

- Rama de la Estadística que pone énfasis en contexto geológico y espacial de los datos.
- Generación de modelos de bloques de leyes 3D para planificación minera.
- Estimación de las reservas locales y globales.
- Cuantificación de incertidumbre en contenido y predicción de la producción.

- Rama de la Estadística que pone énfasis en contexto geológico y espacial de los datos.
- Generación de modelos de bloques de leyes 3D para planificación minera.
- Estimación de las reservas locales y globales.
- Cuantificación de incertidumbre en contenido y predicción de la producción.

Ejemplo: Reservas de Mina Los Bronces, Región Metropolitana



# Kriging

- Dados los puntos  $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_n$  y una variable espacial  $Z$ , conociendo las observaciones  $Z(\mathbf{u}_1), \dots, Z(\mathbf{u}_n)$ , se quiere estimar  $Z(\mathbf{u}_0)$ .
- $Z^*(\mathbf{u}_0) = a + \sum_{i=1}^n \lambda_i Z(\mathbf{u}_i)$
- $Z^*(\mathbf{u}_0)$  debe ser insesgado y de mínima varianza:

$$\mathbb{E}(Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)) = 0$$

$$\frac{\partial}{\partial \lambda_i} \text{Var}(Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)) = 0, \quad i = 1, \dots, n$$

- Distintos tipos: **Simple**, Ordinario, con deriva, no lineal, Cokriging, ...
- La estimación es *suave* ( $\text{Var}(Z^*(\mathbf{u}_0)) < \text{Var}(Z(\mathbf{u}_0))$ ).

# Kriging

- Dados los puntos  $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_n$  y una variable espacial  $Z$ , conociendo las observaciones  $Z(\mathbf{u}_1), \dots, Z(\mathbf{u}_n)$ , se quiere estimar  $Z(\mathbf{u}_0)$ .
- $Z^*(\mathbf{u}_0) = a + \sum_{i=1}^n \lambda_i Z(\mathbf{u}_i)$
- $Z^*(\mathbf{u}_0)$  debe ser insesgado y de mínima varianza:

$$\mathbb{E}(Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)) = 0$$

$$\frac{\partial}{\partial \lambda_i} \text{Var}(Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)) = 0, \quad i = 1, \dots, n$$

- Distintos tipos: **Simple**, Ordinario, con deriva, no lineal, Cokriging, ...
- La estimación es *suave* ( $\text{Var}(Z^*(\mathbf{u}_0)) < \text{Var}(Z(\mathbf{u}_0))$ ).

# Kriging

- Dados los puntos  $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_n$  y una variable espacial  $Z$ , conociendo las observaciones  $Z(\mathbf{u}_1), \dots, Z(\mathbf{u}_n)$ , se quiere estimar  $Z(\mathbf{u}_0)$ .
- $Z^*(\mathbf{u}_0) = a + \sum_{i=1}^n \lambda_i Z(\mathbf{u}_i)$
- $Z^*(\mathbf{u}_0)$  debe ser insesgado y de mínima varianza:

$$\mathbb{E}(Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)) = 0$$

$$\frac{\partial}{\partial \lambda_i} \text{Var}(Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)) = 0, \quad i = 1, \dots, n$$

- Distintos tipos: **Simple**, Ordinario, con deriva, no lineal, Cokriging, ...
- La estimación es *suave* ( $\text{Var}(Z^*(\mathbf{u}_0)) < \text{Var}(Z(\mathbf{u}_0))$ ).



# Kriging

- Dados los puntos  $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_n$  y una variable espacial  $Z$ , conociendo las observaciones  $Z(\mathbf{u}_1), \dots, Z(\mathbf{u}_n)$ , se quiere estimar  $Z(\mathbf{u}_0)$ .
- $Z^*(\mathbf{u}_0) = a + \sum_{i=1}^n \lambda_i Z(\mathbf{u}_i)$
- $Z^*(\mathbf{u}_0)$  debe ser insesgado y de mínima varianza:

$$\mathbb{E}(Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)) = 0$$

$$\frac{\partial}{\partial \lambda_i} \text{Var}(Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)) = 0, \quad i = 1, \dots, n$$

- Distintos tipos: **Simple**, Ordinario, con deriva, no lineal, Cokriging, ...
- La estimación es *suave* ( $\text{Var}(Z^*(\mathbf{u}_0)) < \text{Var}(Z(\mathbf{u}_0))$ ).

# Kriging

- Dados los puntos  $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_n$  y una variable espacial  $Z$ , conociendo las observaciones  $Z(\mathbf{u}_1), \dots, Z(\mathbf{u}_n)$ , se quiere estimar  $Z(\mathbf{u}_0)$ .
- $Z^*(\mathbf{u}_0) = a + \sum_{i=1}^n \lambda_i Z(\mathbf{u}_i)$
- $Z^*(\mathbf{u}_0)$  debe ser insesgado y de mínima varianza:

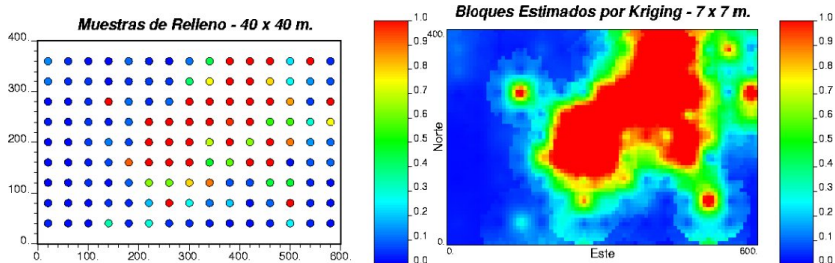
$$\mathbb{E}(Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)) = 0$$

$$\frac{\partial}{\partial \lambda_i} \text{Var}(Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)) = 0, \quad i = 1, \dots, n$$

- Distintos tipos: **Simple**, Ordinario, con deriva, no lineal, Cokriging, ...
- La estimación es *suave* ( $\text{Var}(Z^*(\mathbf{u}_0)) < \text{Var}(Z(\mathbf{u}_0))$ ).

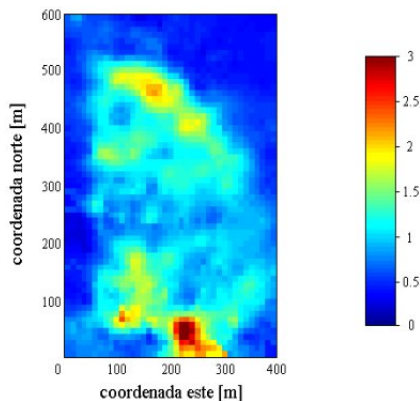
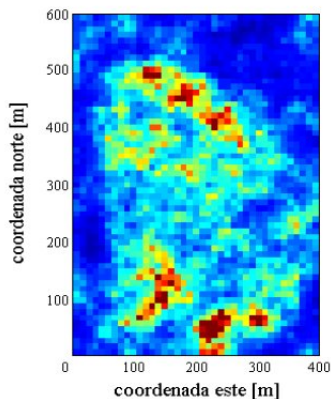
# Kriging

Ejemplo: Estimación vía Kriging sobre una grilla rectangular



# Kriging

Ejemplo: Comparación entre imagen real y estimación vía Kriging



# Simulación Convencional

- Kriging no representa la variabilidad espacial de los datos (*suaviza*).
- Se agrega un residuo aleatorio a la estimación por Kriging:

$$Z_s(\mathbf{u}_0) = Z^*(\mathbf{u}_0) + R(\mathbf{u}_0)$$

- Construyendo varias *realizaciones* se obtienen escenarios posibles para la incerteza.
- Existe una dependencia de la estimación y la simulación con respecto a la **covarianza espacial entre 2 puntos**,  $\text{Cov}(\mathbf{u}_i, \mathbf{u}_j)$  (potencial pérdida de información).

# Simulación Convencional

- Kriging no representa la variabilidad espacial de los datos (*suaviza*).
- Se agrega un residuo aleatorio a la estimación por Kriging:

$$Z_s(\mathbf{u}_0) = Z^*(\mathbf{u}_0) + R(\mathbf{u}_0)$$

- Construyendo varias *realizaciones* se obtienen escenarios posibles para la incerteza.
- Existe una dependencia de la estimación y la simulación con respecto a la **covarianza espacial entre 2 puntos**,  $\text{Cov}(\mathbf{u}_i, \mathbf{u}_j)$  (potencial pérdida de información).

# Simulación Convencional

- Kriging no representa la variabilidad espacial de los datos (*suaviza*).
- Se agrega un residuo aleatorio a la estimación por Kriging:

$$Z_s(\mathbf{u}_0) = Z^*(\mathbf{u}_0) + R(\mathbf{u}_0)$$

- Construyendo varias *realizaciones* se obtienen escenarios posibles para la incerteza.
- Existe una dependencia de la estimación y la simulación con respecto a la **covarianza espacial entre 2 puntos**,  $\text{Cov}(\mathbf{u}_i, \mathbf{u}_j)$  (potencial pérdida de información).

# Simulación Convencional

- Kriging no representa la variabilidad espacial de los datos (*suaviza*).
- Se agrega un residuo aleatorio a la estimación por Kriging:

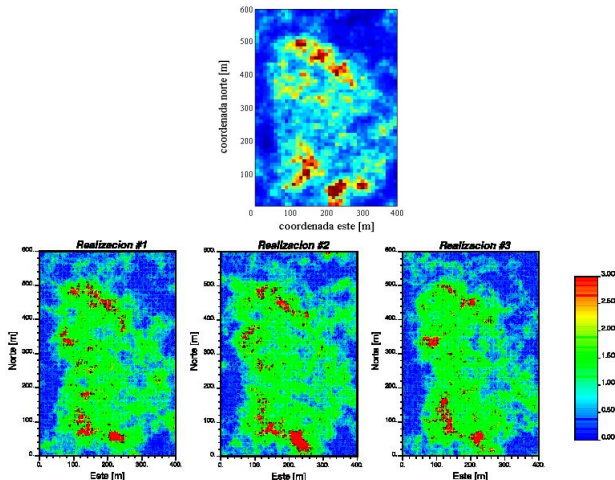
$$Z_s(\mathbf{u}_0) = Z^*(\mathbf{u}_0) + R(\mathbf{u}_0)$$

- Construyendo varias *realizaciones* se obtienen escenarios posibles para la incerteza.
- Existe una dependencia de la estimación y la simulación con respecto a la **covarianza espacial entre 2 puntos**,  $\text{Cov}(\mathbf{u}_i, \mathbf{u}_j)$  (potencial pérdida de información).



# Simulación Convencional

Ejemplo: Distintas realizaciones obtenidas con Simulación Convencional



# Simulación no Convencional: Recocido Simulado

- Heurística proveniente de la Optimización Combinatorial para resolver problemas *grandes* mediante exploración de estados o configuraciones.
- Entrega *buenos* óptimos globales.
- Vendedor viajero, ruteo de vehículos, scheduling, layout, coloreo de grafos, asignación cuadrática, bin packing,...

# Simulación no Convencional: Recocido Simulado

- Heurística proveniente de la Optimización Combinatorial para resolver problemas *grandes* mediante exploración de estados o configuraciones.
- Entrega *buenos* óptimos globales.
- Vendedor viajero, ruteo de vehículos, scheduling, layout, coloreo de grafos, asignación cuadrática, bin packing,...

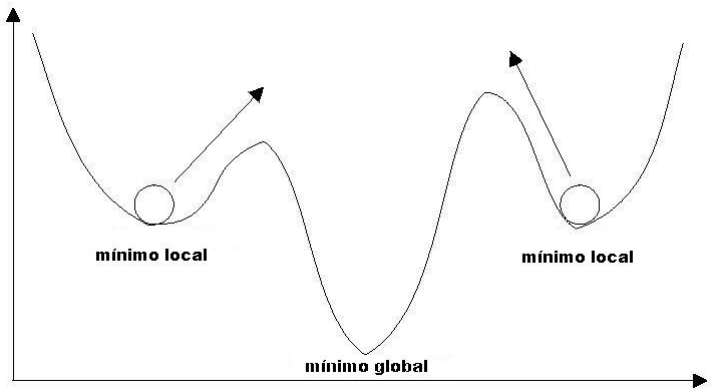
# Simulación no Convencional: Recocido Simulado

- Heurística proveniente de la Optimización Combinatorial para resolver problemas *grandes* mediante exploración de estados o configuraciones.
- Entrega *buenos* óptimos globales.
- Vendedor viajero, ruteo de vehículos, scheduling, layout, coloreo de grafos, asignación cuadrática, bin packing,...

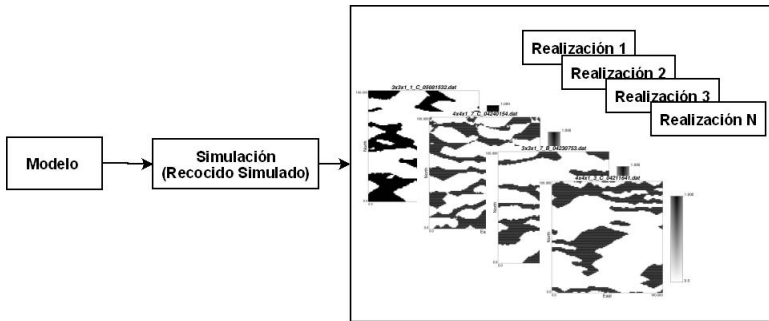
# Simulación no Convencional: Recocido Simulado

$$\mathbb{P}(\text{aceptar estado } k \text{ a partir de estado } i) = \begin{cases} 1 & O_k \leq O_i \\ e^{(O_i - O_k)/T} & O_k > O_i \end{cases}$$

## Simulación no Convencional: Recocido Simulado



# Simulación no Convencional: Recocido Simulado

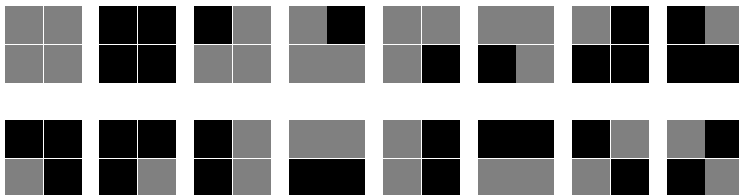


# Patrones 2D

- Estadísticas de Múltiples Puntos: Patrones 2D
- Template



- Patrones



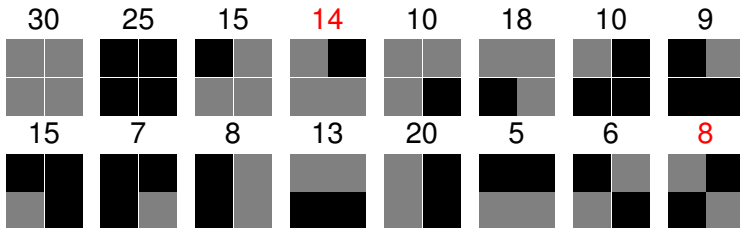


# Patrones 2D

- Estadísticas de Múltiples Puntos: Patrones 2D
- Template



- Histograma de frecuencias de Patrones

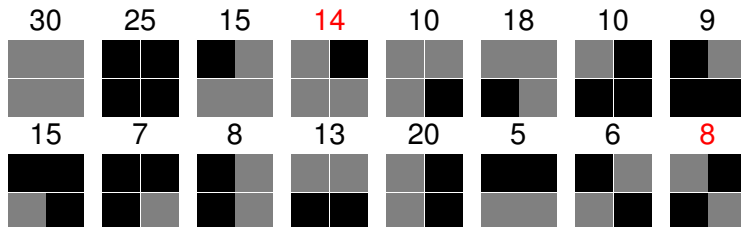


$$\text{Frecuencia de Patrones} \approx \underbrace{\mathbb{P}(Z(\mathbf{u}_0) = z | Z(\mathbf{u}_1) = z_1, \dots, Z(\mathbf{u}_n) = z_n)}_{\text{Probabilidad Condicional}}$$

# Patrones 2D

Ejemplo:

$u_1$	$u_3$
$u_0$	$u_2$



# Patrones 2D

Ejemplo:

$u_1$	$u_3$
$u_0$	$u_2$

14



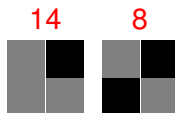
8



# Patrones 2D

Ejemplo:

$u_1$	$u_3$
$u_0$	$u_2$

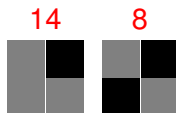


**¿Cuál es la probabilidad de que  $Z(u_0)$  sea negro/gris dado que  $Z(u_1)$  =gris,  $Z(u_2)$  =gris y  $Z(u_3)$  =negro?**

# Patrones 2D

Ejemplo:

$\mathbf{u}_1$	$\mathbf{u}_3$
$\mathbf{u}_0$	$\mathbf{u}_2$

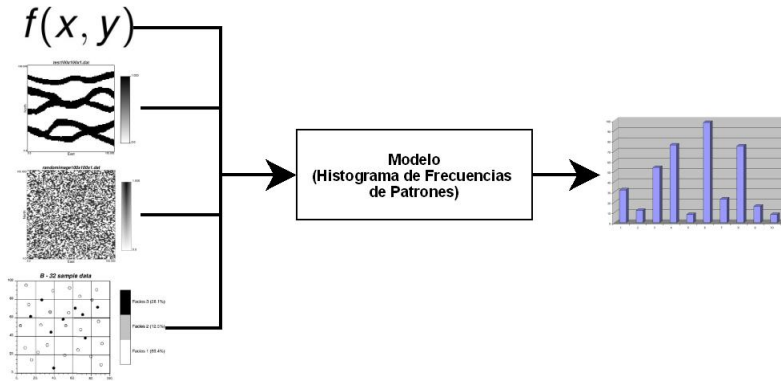


**¿Cuál es la probabilidad de que  $Z(\mathbf{u}_0)$  sea negro/gris dado que  $Z(\mathbf{u}_1) = \text{gris}$ ,  $Z(\mathbf{u}_2) = \text{gris}$  y  $Z(\mathbf{u}_3) = \text{negro}$ ?**

$$\mathbb{P}(Z(\mathbf{u}_0) = \text{negro} \mid Z(\mathbf{u}_1) = \text{gris}, Z(\mathbf{u}_2) = \text{gris}, Z(\mathbf{u}_3) = \text{negro}) = \frac{8}{8 + 14}$$

$$\mathbb{P}(Z(\mathbf{u}_0) = \text{gris} \mid Z(\mathbf{u}_1) = \text{gris}, Z(\mathbf{u}_2) = \text{gris}, Z(\mathbf{u}_3) = \text{negro}) = \frac{14}{8 + 14}$$

# Patrones 2D

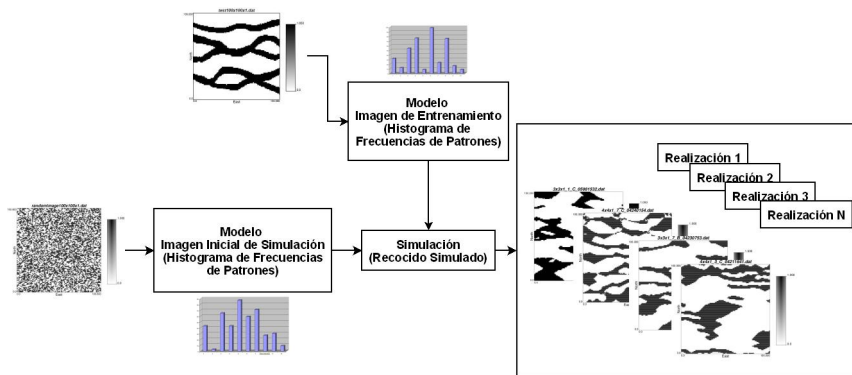


# Esquema

- 1 Introducción
- 2 Antecedentes
- 3 Trabajo realizado**
- 4 Resultados
- 5 Conclusiones y Trabajo a futuro



# Diagrama de la simulación



# Principales tareas

- Contrucción de histograma de frecuencias de patrones para una imagen
- Implementar Recocido Simulado para obtener imágenes simuladas a partir de imagen de entrenamiento

# Principales tareas

- Contrucción de histograma de frecuencias de patrones para una imagen
- Implementar Recocido Simulado para obtener imágenes simuladas a partir de imagen de entrenamiento

# Principales problemas

- Manejo de patrones grandes
- Tiempo de cálculo del Recocido Simulado

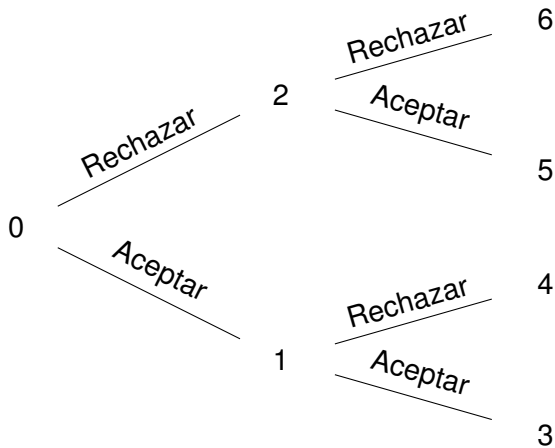
# Principales problemas

- Manejo de patrones grandes
- Tiempo de cálculo del Recocido Simulado

- Manejo de patrones grandes: se optó por realizar pruebas hasta patrones de  $4 \times 4$  (requiere mayor investigación)
- Tiempo de cálculo del Recocido Simulado: *Paralelización*

- Manejo de patrones grandes: se optó por realizar pruebas hasta patrones de  $4 \times 4$  (requiere mayor investigación)
- Tiempo de cálculo del Recocido Simulado: ***Paralelización***

# Computación Especulativa para Recocido Simulado





# Herramientas utilizadas

- C++: `map<string, int>`
- MPI (*Message Passing Interface*)
- GSLIB: `pixelplt`

# Herramientas utilizadas

- C++: `map<string, int>`
- MPI (*Message Passing Interface*)
- GSLIB: `pixelplt`

# Herramientas utilizadas

- C++: `map<string, int>`
- MPI (*Message Passing Interface*)
- GSLIB: `pixelplt`

# Speedup teórico

- Utilizando  $P$  procesos, se obtiene speedup  $\log_2(P + 1)$
- Para 3 procesos, se obtiene speedup de 2 (reducción a la mitad del tiempo)
- Para 7 procesos, se obtiene speedup de 3 (reducción a un tercio del tiempo)

# Speedup teórico

- Utilizando  $P$  procesos, se obtiene speedup  $\log_2(P + 1)$
- Para 3 procesos, se obtiene speedup de 2 (reducción a la mitad del tiempo)
- Para 7 procesos, se obtiene speedup de 3 (reducción a un tercio del tiempo)

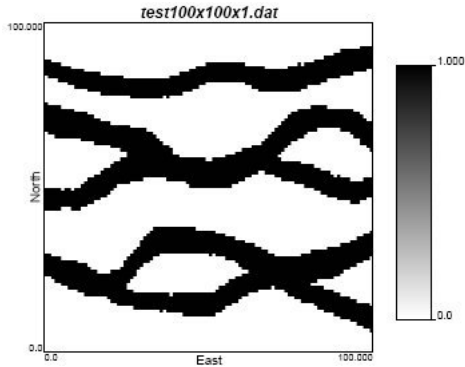
# Speedup teórico

- Utilizando  $P$  procesos, se obtiene speedup  $\log_2(P + 1)$
- Para 3 procesos, se obtiene speedup de 2 (reducción a la mitad del tiempo)
- Para 7 procesos, se obtiene speedup de 3 (reducción a un tercio del tiempo)

# Esquema

- 1 Introducción
- 2 Antecedentes
- 3 Trabajo realizado
- 4 Resultados**
- 5 Conclusiones y Trabajo a futuro

# Imagen de entrenamiento





# Funciones Objetivo

- **Sin pesos** asociados a las frecuencias:

$$O = \sum_{i \in \mathcal{P}} (f_i^{TI} - f_i^{RE})^2$$

- **Con pesos** asociados a las frecuencias:

$$O = \sum_{i \in \mathcal{P}} \left\{ \frac{\frac{1}{f_i^{TI}}}{\sum_{j \in \mathcal{P}} \frac{1}{f_j^{TI}}} \right\} (f_i^{TI} - f_i^{RE})^2$$

donde  $\frac{1}{f_i^{TI}} = 0$ , si  $f_i^{TI} = 0$

# Funciones Objetivo

- **Sin pesos** asociados a las frecuencias:

$$O = \sum_{i \in \mathcal{P}} (f_i^{TI} - f_i^{RE})^2$$

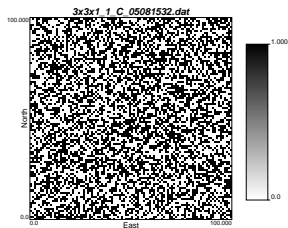
- **Con pesos** asociados a las frecuencias:

$$O = \sum_{i \in \mathcal{P}} \left\{ \frac{\frac{1}{f_i^{TI}}}{\sum_{j \in \mathcal{P}} \frac{1}{f_j^{TI}}} \right\} (f_i^{TI} - f_i^{RE})^2$$

donde  $\frac{1}{f_i^{TI}} = 0$ , si  $f_i^{TI} = 0$

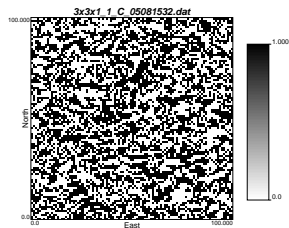
# Resultados: Evolución de una Simulación

Iteración 0



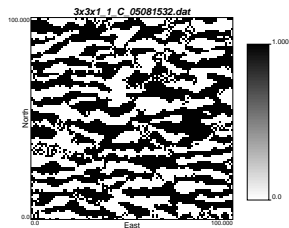
# Resultados: Evolución de una Simulación

Iteración 349860



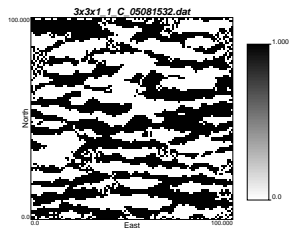
# Resultados: Evolución de una Simulación

Iteración 529788



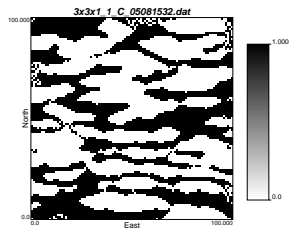
# Resultados: Evolución de una Simulación

Iteración 699720



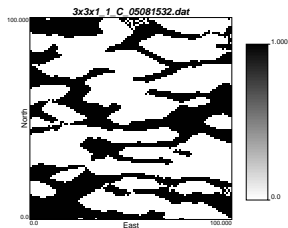
# Resultados: Evolución de una Simulación

Iteración 879648



# Resultados: Evolución de una Simulación

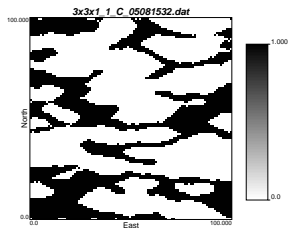
Iteración 1049580





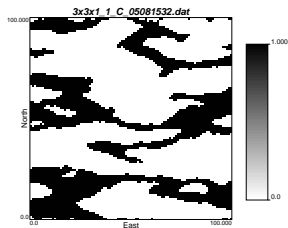
# Resultados: Evolución de una Simulación

Iteración 1229508



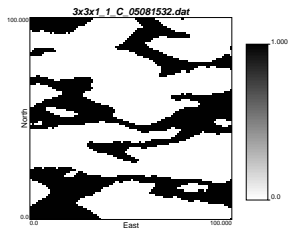
# Resultados: Evolución de una Simulación

Iteración 1579368



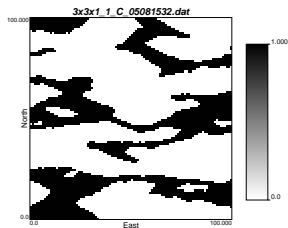
# Resultados: Evolución de una Simulación

Iteración 1749300



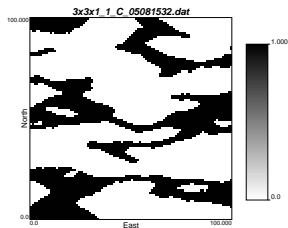
# Resultados: Evolución de una Simulación

Iteración 1929228



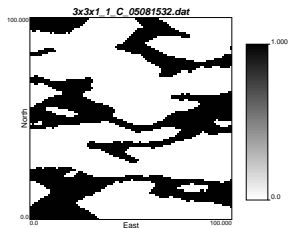
# Resultados: Evolución de una Simulación

Iteración 2059176












# Resultados: Evolución de una Simulación










Iteración 2109156



# Resultados: Imágenes Simuladas sin pesos

	Escenario A	Escenario B	Escenario C
$2 \times 2 \times 1$			
$3 \times 3 \times 1$			
$4 \times 4 \times 1$			

# Resultados: Imágenes Simuladas con pesos

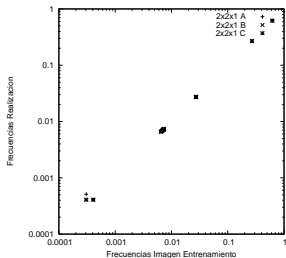
	Escenario A	Escenario B	Escenario C
$2 \times 2 \times 1$			
$3 \times 3 \times 1$			
$4 \times 4 \times 1$			



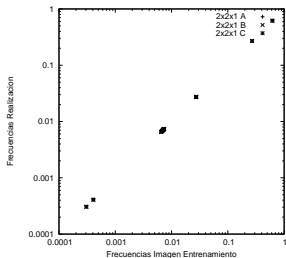
# Resultados: Comparación de Histogramas

$2 \times 2 \times 1$

Sin pesos



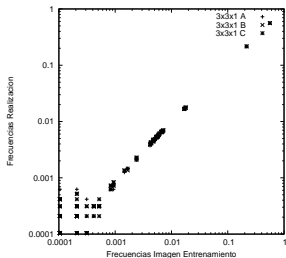
Con pesos



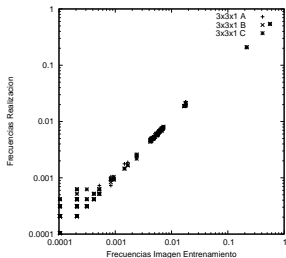
# Resultados: Comparación de Histogramas

$3 \times 3 \times 1$

Sin pesos



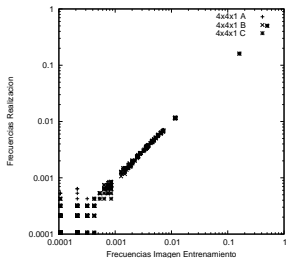
Con pesos



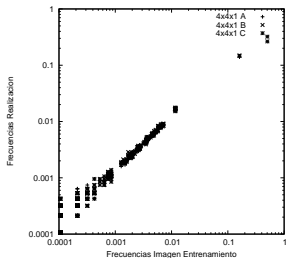
# Resultados: Comparación de Histogramas

$4 \times 4 \times 1$

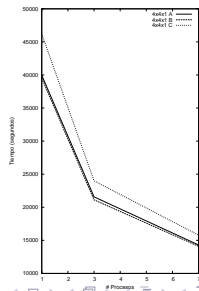
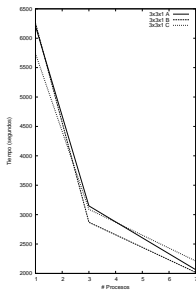
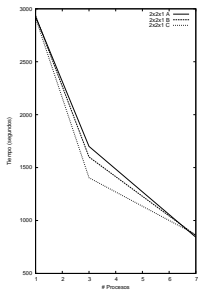
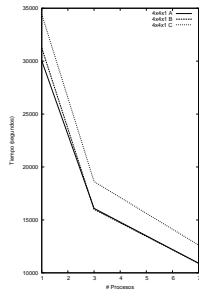
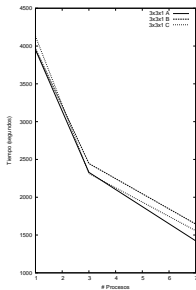
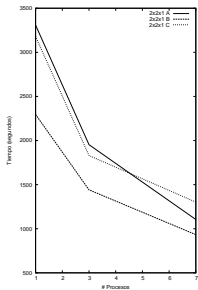
Sin pesos



Con pesos



# Resultados: Speedup Experimental



# Esquema

- 1 Introducción
- 2 Antecedentes
- 3 Trabajo realizado
- 4 Resultados
- 5 Conclusiones y Trabajo a futuro**

# Conclusiones

- Reducción de tiempo de cálculo del Recocido Simulado con un speedup de  $\log_2(P + 1)$ .

# Conclusiones

- Reducción de tiempo de cálculo del Recocido Simulado con un speedup de  $\log_2(P + 1)$ .
- Utilización de pesos en función objetivo influye levemente en la obtención de mejores realizaciones.

# Conclusiones

- Reducción de tiempo de cálculo del Recocido Simulado con un speedup de  $\log_2(P + 1)$ .
- Utilización de pesos en función objetivo influye levemente en la obtención de mejores realizaciones.
- Trade-off: mayor tiempo de cálculo a medida que crece el tamaño del template.



# Conclusiones

- Reducción de tiempo de cálculo del Recocido Simulado con un speedup de  $\log_2(P + 1)$ .
- Utilización de pesos en función objetivo influye levemente en la obtención de mejores realizaciones.
- Trade-off: mayor tiempo de cálculo a medida que crece el tamaño del template.
- Speedup teórico  $\log_2(P + 1)$  es alcanzado con mayor precisión cuando se utiliza una función objetivo con pesos.

# Trabajo a futuro

- Optimización de las estructuras de datos utilizadas para manejar los patrones.

# Trabajo a futuro

- Optimización de las estructuras de datos utilizadas para manejar los patrones.
- Incorporación de árboles no balanceados en Computación Especulativa para Recocido Simulado.

# Trabajo a futuro

- Optimización de las estructuras de datos utilizadas para manejar los patrones.
- Incorporación de árboles no balanceados en Computación Especulativa para Recocido Simulado.
- Utilización de templates no regulares.

# Trabajo a futuro

- Optimización de las estructuras de datos utilizadas para manejar los patrones.
- Incorporación de árboles no balanceados en Computación Especulativa para Recocido Simulado.
- Utilización de templates no regulares.
- Realización de pruebas utilizando una mayor cantidad de procesos (15, 31, 63,...).

¿Preguntas?