Presentación del proyecto final

Un análisis geoestadístico se juzga por las decisiones que permite tomar. En esta lección conviertes el notebook del proyecto en un reporte técnico breve que una autoridad del agua pueda usar para decidir dónde se puede bombear y dónde perforar el próximo pozo de monitoreo. Este enunciado, con la rúbrica, también se adjunta en PDF para que lo tengas a mano mientras trabajas.

El encargo

Tienes 150 pozos con la profundidad del nivel freático (metros bajo la superficie) en una zona de 50 × 50 km, generados con `generar_pozos(semilla=42)`. Bombear deja de ser rentable cuando el agua está a más de $u = 32$ m. Debes entregar:

1. Un mapa de la profundidad estimada en una malla de 1 km.
2. Un mapa de la incertidumbre (desviación estándar de kriging).
3. Un mapa de la probabilidad de que la profundidad supere 32 m y el porcentaje del área donde esa probabilidad es mayor que 0.8.
4. Una recomendación de ubicación para un nuevo pozo de monitoreo, justificada con el mapa de incertidumbre.

Estructura del reporte

1. **Datos y exploración** (1 párrafo + 1 figura): distribución de profundidades, mapa de puntos y la tendencia detectada con sus coeficientes.
2. **Variograma** (1 figura): experimental de los datos crudos y de los residuos, con el modelo elegido y sus parámetros. Explica qué cambia al quitar la tendencia.
3. **Comparación de modelos** (1 tabla): RMSE, error medio y MSDR de validación cruzada para la media, IDW, kriging ordinario y kriging universal.
4. **Mapas** (3 paneles): estimación, desviación estándar y probabilidad de superar el umbral.
5. **Recomendación** (1 párrafo): dónde se puede bombear con confianza, dónde no y dónde hace falta más información.
6. **Limitaciones** (3-5 viñetas): supuestos de estacionariedad de los residuos, normalidad para el mapa de probabilidad, efecto borde y variograma no reajustado en cada pliegue.

Calcular la fracción del área

Con la malla de estimación y su varianza, la probabilidad de excedencia y la fracción de área salen en tres líneas:

```python
from scipy.stats import norm

prob = 1 - norm.cdf((32 - z_est) / np.sqrt(varianza))
fraccion = (prob > 0.8).mean()
print(f"{100 * fraccion:.1f}% del área con P(z > 32 m) > 0.8")
```

Como la malla es regular, cada celda pesa lo mismo y la media de la máscara booleana es directamente la fracción del área.

Rúbrica (100 puntos)

- **Variograma (25):** experimental bien construido (corte en la mitad de la distancia máxima, pares por lag reportados) y modelo justificado.
- **Validación (25):** LOO para al menos cuatro métodos, con RMSE, ME y MSDR; comparación contra la media.
- **Mapas (20):** ejes en km, barras de color con unidades, pozos superpuestos en el mapa de incertidumbre.
- **Interpretación (20):** la recomendación se sigue de los mapas y reconoce la incertidumbre.
- **Reproducibilidad (10):** el notebook corre de principio a fin y reproduce los números del reporte.

Extensiones opcionales

- Repite el análisis con `semilla=7` y compara qué tan estable es el variograma.
- Prueba `exact_values=False` con una pepita de 0.25 y discute cómo cambia el mapa junto a los pozos.
- Usa `GridSearchCV` con `pykrige.rk.Krige` y explica por qué su resultado puede diferir de tu LOO.

Trampas comunes

- Presentar el mapa de estimación sin el de incertidumbre.
- Recomendar bombeo en zonas sin pozos solo porque el mapa estimado muestra valores bajos.
- Reportar un RMSE sin decir con qué tipo de validación se obtuvo.

Cierre

Un buen reporte geoestadístico cabe en dos páginas: un variograma, una tabla y tres mapas. Si cada figura responde una pregunta del encargo, el trabajo está completo.