Pacific Northwest
NATIONAL LABORATORY

Organized in cooperation with

IAEA
International Atomic Energy Agency
Atoms for Peace and Development

Principal Component Analysis (PCA) to compute major directions for 3D spatial analysis

November 15, 2023

Swasti Saxena
Post Doctorate Research Associate

Contributing authors: Moses Obiri, Deborah Fagan, Jenniffer Huckett

2023 Global Summit
on Environmental Remediation
@REMPLEX

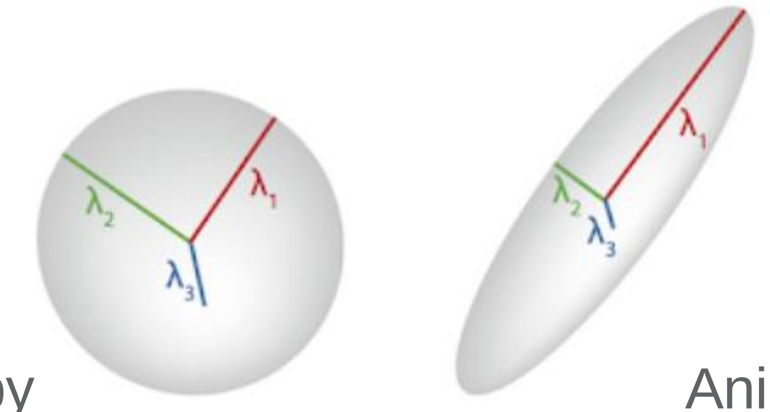
U.S. DEPARTMENT OF ENERGY
BATTELLE

PNNL is operated by Battelle for the U.S. Department of Energy

Introduction

- Accurate spatial modeling is crucial for predicting contaminant spread for subsurface environmental clean-up.

- 3D anisotropic models account for variation in a property in three principal directions.



- Traditionally, these directions are chosen through visual inspection of variograms.

- We present an automated method to objectively identify these three principal directions of anisotropy in 3D spatial data.

Principal Component Analysis (PCA)

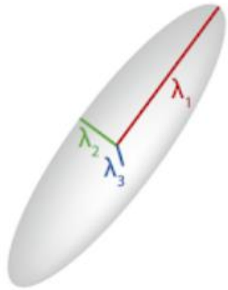
- Uses linear algebra to reduce dimensionality or extract prominent features.
- A covariance matrix is computed to establish the spread of variances between different variables/features.
- Eigenvalue decomposition of the covariance matrix → eigenvector matrix, which would rotate the coordinate system along the principal component.
- Moment of Inertia (MOI) method was explored to establish the spread of variance between the variable and its ‘2D/3D location’

Moment of Inertia (MOI) Method

- An MOI tensor is the sum of the mass distribution occurring in a rigid body rotating about the axes of rotation.
- Calculated by multiplying mass with its distance from the axis of rotation

$$I_{xx} = \sum_i m_i(y_i^2 + z_i^2)$$

- In spatial analysis, the mass is replaced by covariance volume.
- Smallest MOI → Principal direction of the MOI.



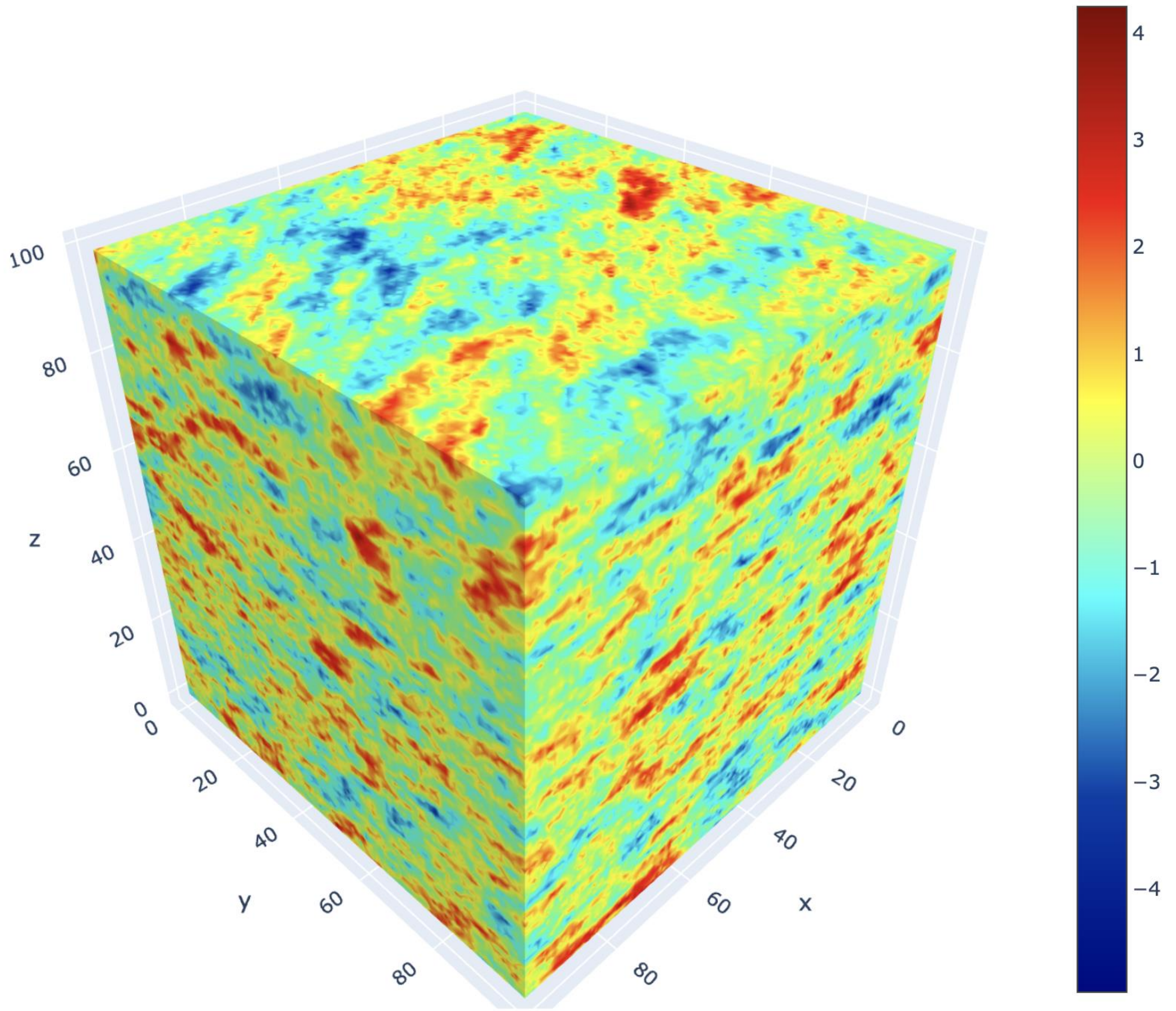
- Eigenvalue decomposition of the covariance matrix → eigenvector matrix, which would rotate the coordinate system along the principal component.

2023 Global Summit

Spatial Data Generated Using GSTools

- GSTools* is a Python library used to create 2D and 3D spatial random field with defined anisotropy in a variogram model.

- Angles of 3D anisotropy are α, β, γ.

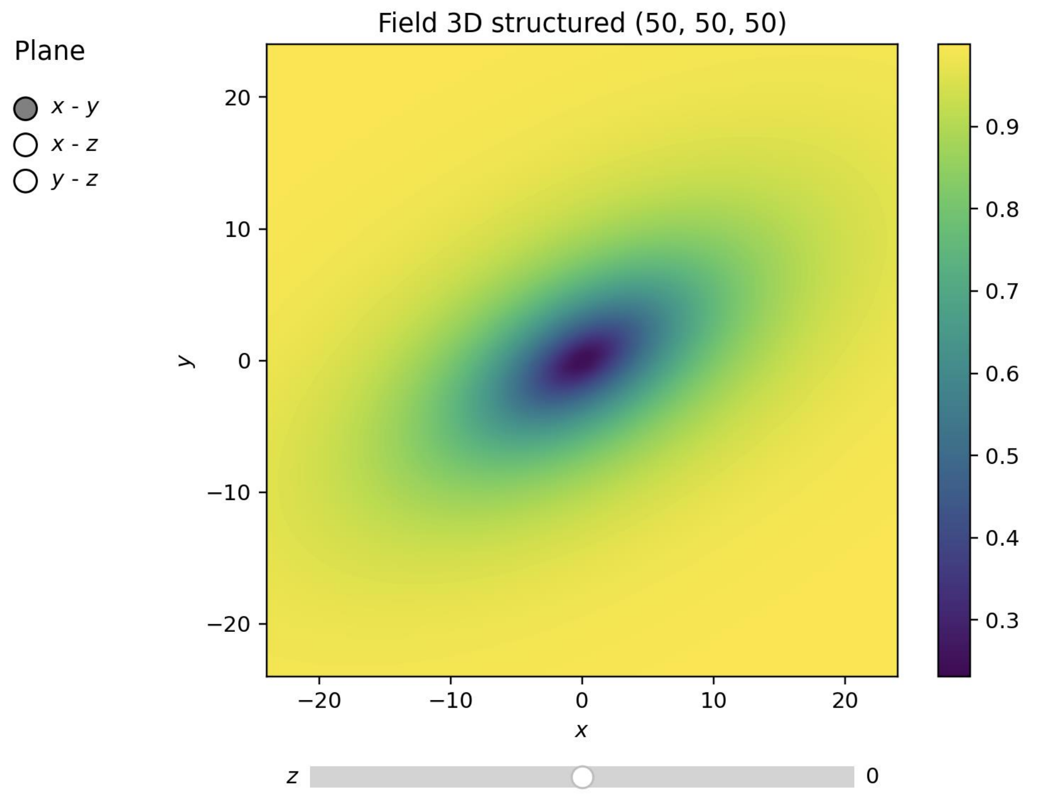


*Müller et al., 2022. https://doi.org/10.5194/gmd-15-3161-2022

2023 Global Summit

Angles of Anisotropy in 2D and 3D

- GSTools follows the Tait-Bryan convention of angular rotations.
- Sequential rotations are made about Z, Y, and X axes, respectively.
- The first rotation about the Z-axis (X-Y plane), α, is the first major direction of anisotropy.
- In 3D anisotropy, β and γ are angular rotations about Y and X axes, respectively.



Metrics Computed to Ascertain Accuracy of Angular Estimates

- We calculate the following metrics to ascertain the accuracy of estimates of α, β, and γ -

  - Mean, $\mu = \frac{\sum_{i=1}^{n} x_i}{n}$

  where n is the total number of estimates (realizations) and x_i are individual predicted values.

  - Mean absolute deviation, MAD $= \frac{\sum_{i=1}^{n} \lfloor x_i - \hat{x} \rfloor}{n}$

  where $\hat{x}$ are the true values.

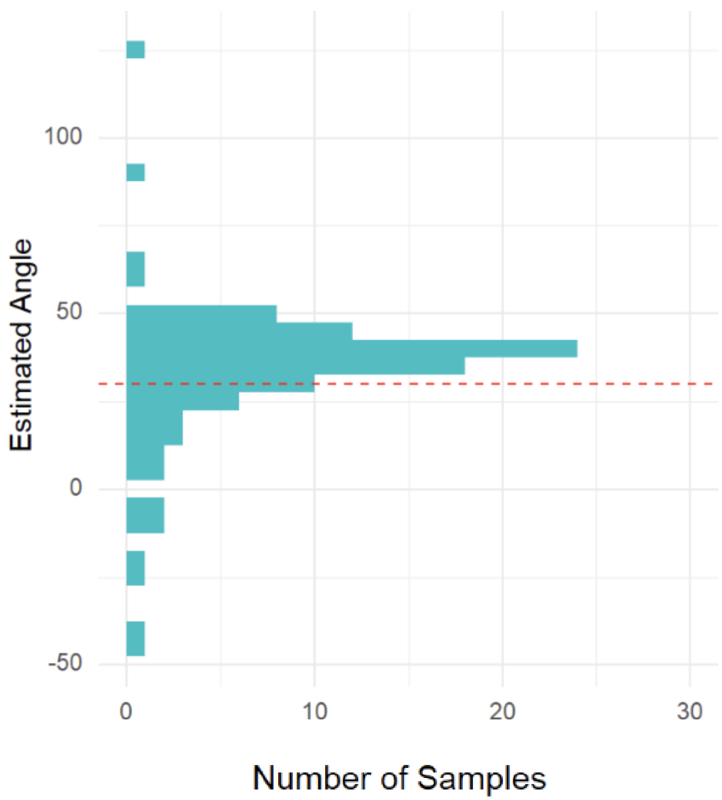
MOI Applied to 2D Dataset

- One hundred realizations of spatial random field are generated based on the following variogram model:
  - model type: Exponential
  - dimensions: 2
  - nugget: 0.0
  - variance: 1.0
  - length scales: 8, 4
  - anisotropy: 0.5

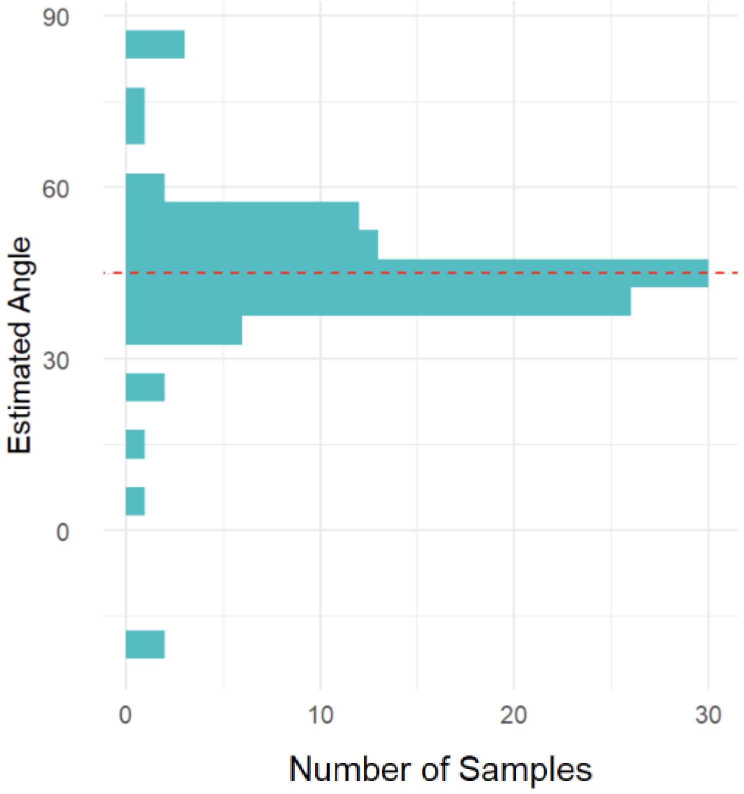
- Angle of anisotropy (α) tested were 30°, 45°, 60°.

- MOI method was applied to each dataset to estimate α.

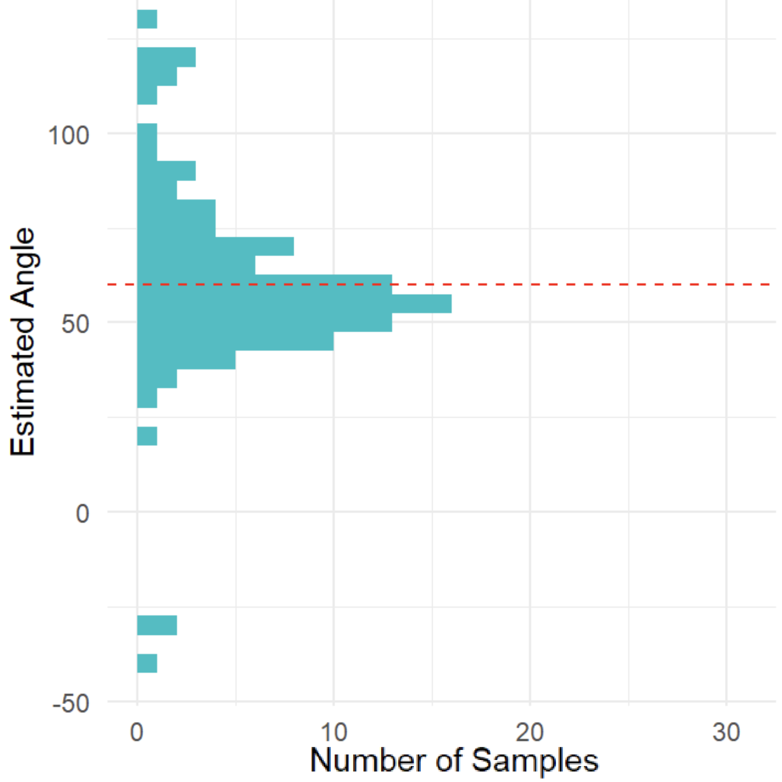
Results for MOI in 2D



α = 30°, Mean = 32.9°, MAD = 9.1°



α = 45°, Mean = 44.3°, MAD = 4.3°

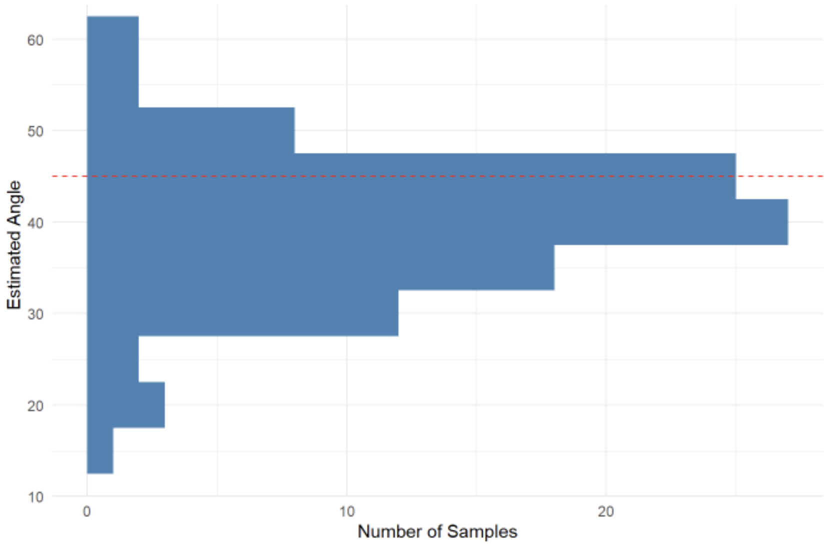


α = 60°, Mean = 59.9°, MAD = 8.6°

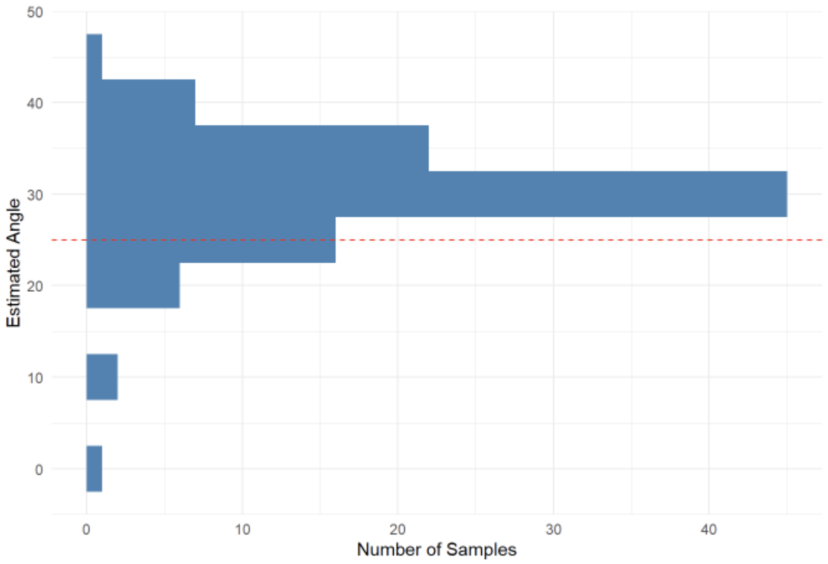
MOI Applied to 3D Dataset

- One hundred realizations of the same variogram model as 2D with the following modifications:
  - length scales: 8, 4, 2
  - anisotropy: 0.5, 0.25
  - angles of anisotropy: α=45°, β=25°, γ=10°

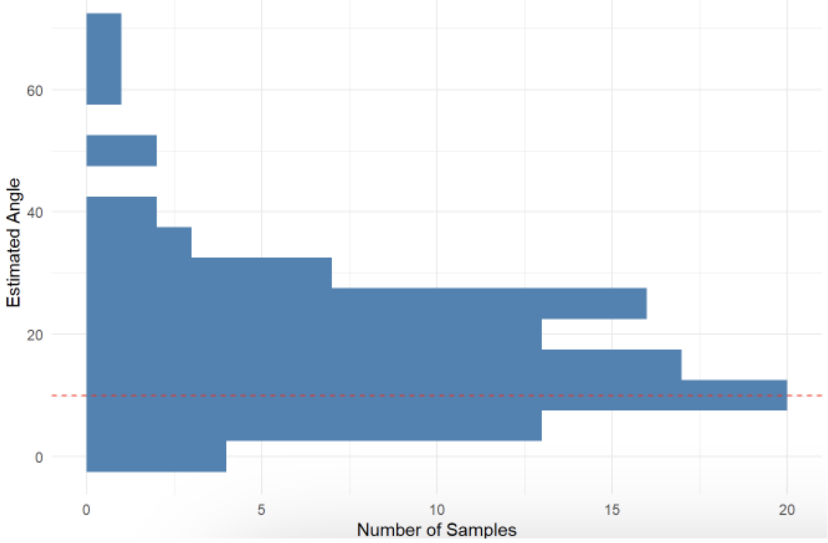
α = 45°, Mean = 39.5°, MAD = 7.4°



β = 25°, Mean = 29.9°, MAD = 6.6°



γ = 10°, Mean = 18.7°, MAD = 10.8°



Conclusions

- MOI method is successful in automating the determination of major angles of anisotropy from a randomly generated spatial field.
- For 2D datasets, the estimated angles are within $\pm 3°$ range and variability is ~4° - 9°.
- For a 3D dataset, the estimated angles are within $\pm$ 9° range and variability is ~6° - 11°.
- We are looking into another approach using the Covariance Tensor Identity (CTI) method to ascertain more accurate automated estimates.
  - In CTI, the covariance function's Hessian matrix is derived from sample derivatives. Anisotropy parameters are determined through solutions of nonlinear equations connecting anisotropic angles with the ratios of the Hessian matrix elements.
- These methods will be applied to real databases to evaluate their practical applications to existing spatial analysis software like the Visual Sample Plan (VSP).

Organized in cooperation with

Questions?

2023 Global Summit
on Environmental Remediation
@REMPLEX

PNNL is operated by Battelle for the U.S. Department of Energy