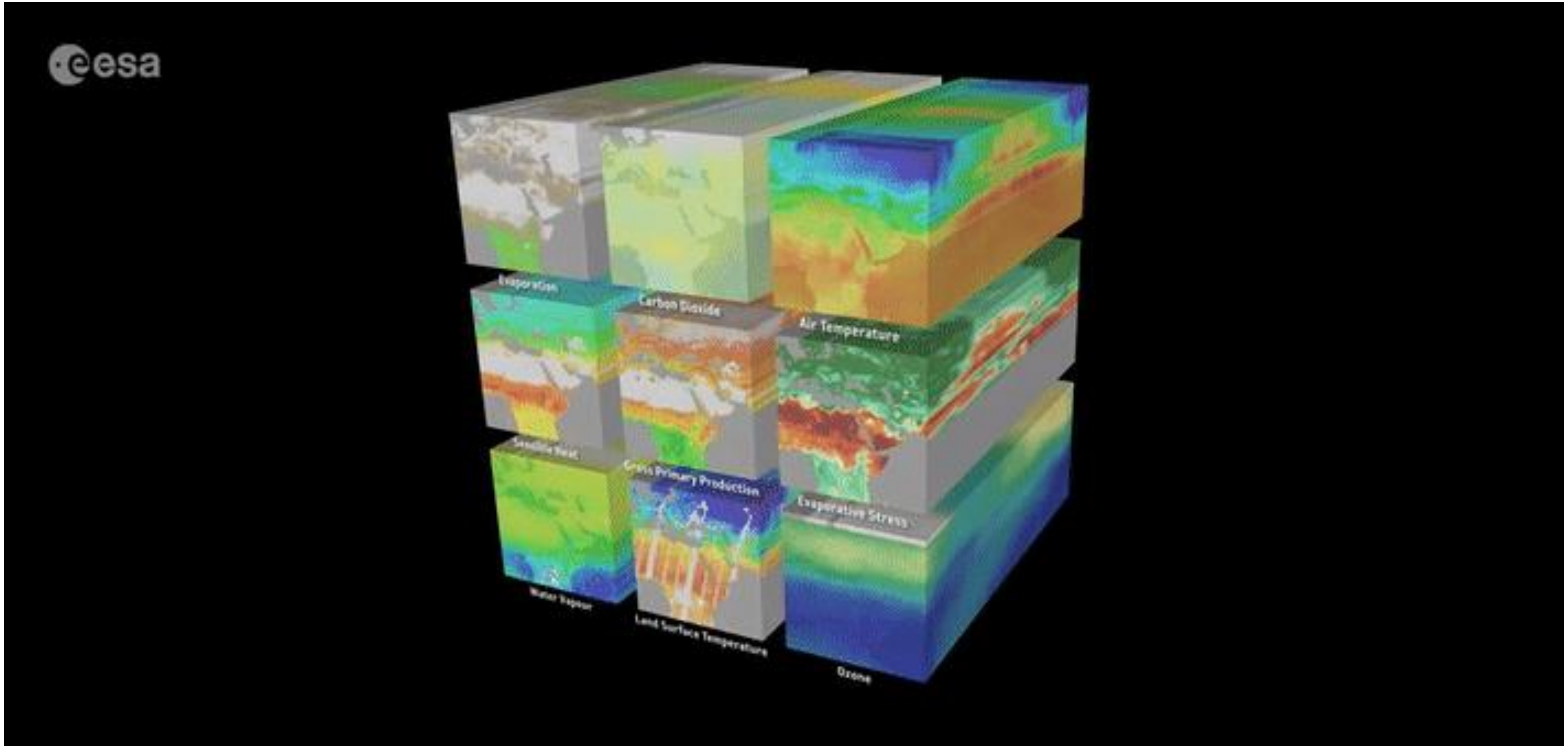
Geospatial tech in your life?



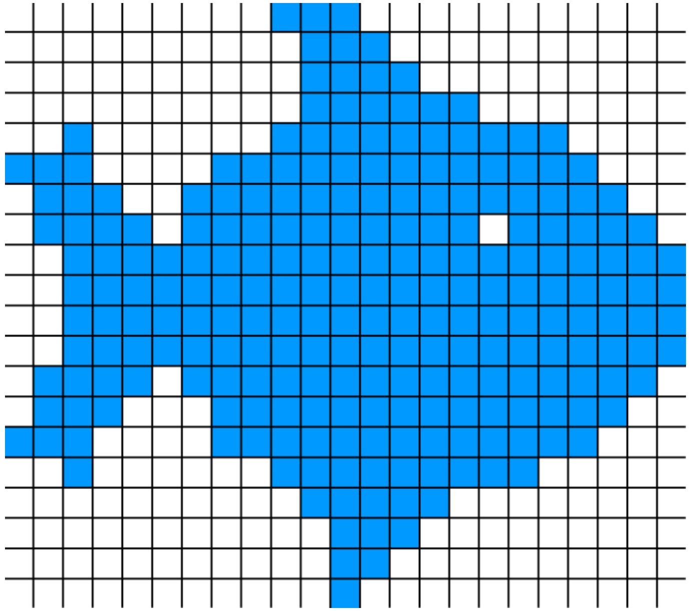
GEOG 450/650 **Geospatial Data**

Instructor: Alex Bevington
alex.bevington@unbc.ca

2025-01-09

Raster

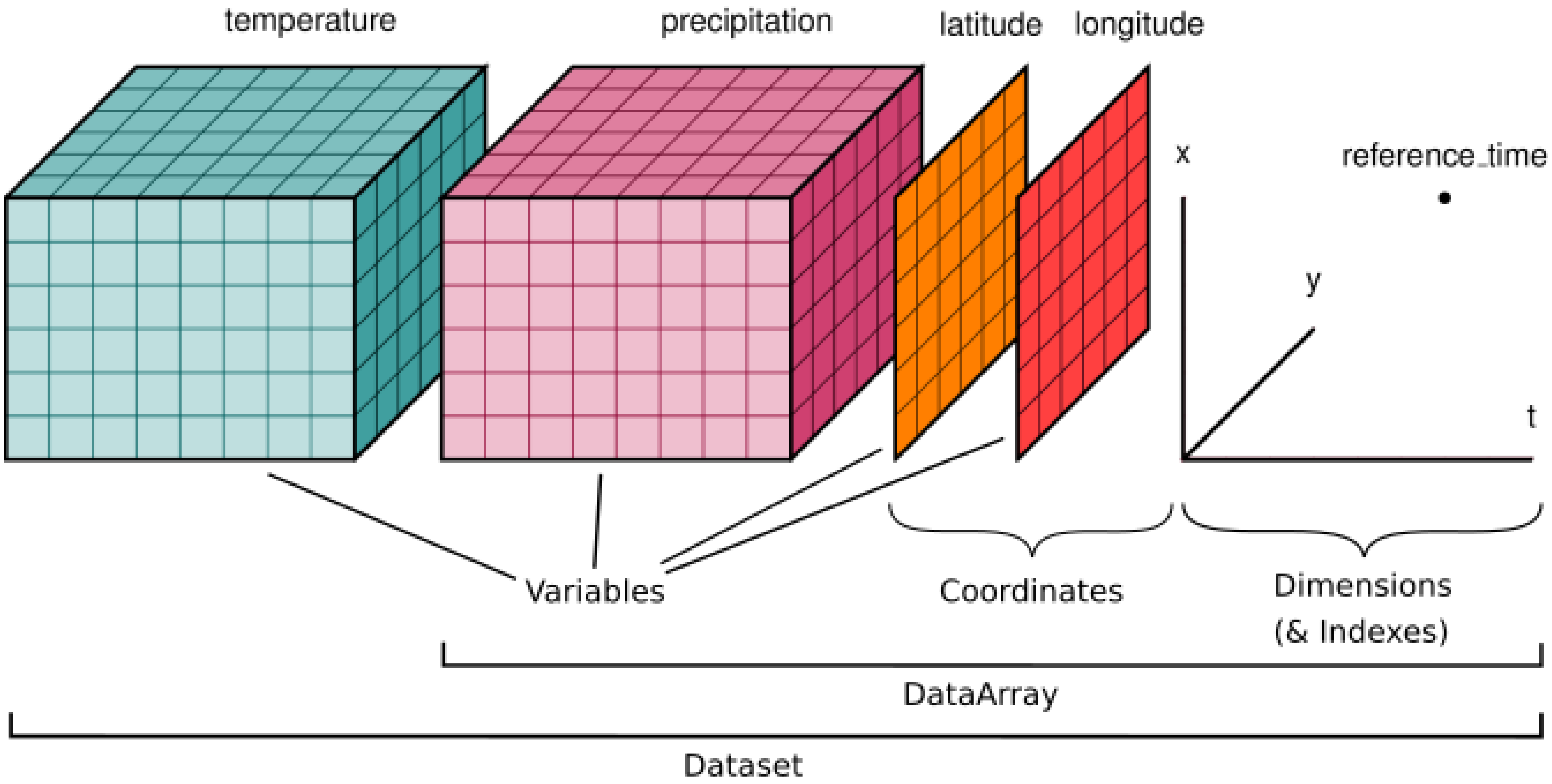
- Grid structure (matrix) of rows and columns
- Defined in map units (e.g., meters, degrees).
- Uses:
  - **Imagery** *multispectral, hyperspectral, thermal*
  - **Radar** *backscatter, phase*
  - **Elevation** *photogrammetry, lidar, interferometry*
  - **Climate** *Interpolation of observations, model outputs*
  - **Classification** *categorical data*



Raster Formats

Feature	GeoTIFF	ASCII	NetCDF
File Type	*Raster (binary)*	*Text (human-readable)*	*Binary (hierarchical)*
Metadata	*Strong (embedded)*	*Weak (basic headers)*	*Strong (self-describing)*
Efficiency	*High*	*Low (large file size)*	*High*
Dimensionality	*Limited*	*Limited*	*Strong (e.g., 3D, 4D)*
Use Cases	*GIS, remote sensing*	*Simple data sharing*	*Scientific analysis*
Tool Support	*GIS tools (QGIS, ArcGIS)*	*Universal (text editors)*	*Scientific libraries*

NetCDF example of storing 4 dimensional data x, y, time



https://docs.xarray.dev/en/stable/user-guide/data-structures.html

Vector

- Data represented by points, lines, or polygons (and MULTI)
- Geometry and attribute data.
- Associated with a CRS (Coordinate Reference System).
- Applications: Road networks, land parcels, boundaries.

Feature	KML/KMZ	GeoJSON	SHP	SQLite	GPKG	GeoParquet	Esri Geodatabase	LAS/LAZ
Type	*Vector (styled)*	*Vector*	*Vector*	*Vector*	*Vector & Raster*	*Vector*	*Vector, Raster, …*	*Point Cloud (3D)*
Structure	*Single XML file (KMZ is zipped)*	*Single JSON-based file*	*Multi-file (shp, shx, dbf, prj)*	*Single SQLite-based file*	*Single SQLite-based file*	*Single binary file*	*Directory (file-based) or database (enterprise)*	*Single binary file (LAS) or compressed (LAZ)*
Efficiency	*Low*	*Low*	*Moderate*	*Moderate*	***High***	***High***	***High***	***High***
Size limit	*Moderate 4GB*	*Moderate ~2GB*	*Moderate 2GB*	***Large ~100TB***	***Large ~100TB***	***Very large ~1PB***	***Large 1TB***	*Moderate ~100GB*
Performance	*Low*	*Low*	***Moderate***	***Moderate***	***High***	***High***	***High***	***High***
Best Use	*Visualization*	*Web maps*	*GIS basics*	*Lightweight DB*	*Mixed datasets*	*Big data/cloud*	*Enterprise GIS*	*LiDAR and 3D data*
Open	***Open***	***Open***	***Open***	***Open***	***Open***	***Open***	*Commercial*	***Open***
Querying	*No*	*No*	*No*	***Yes***	***Yes***	***Yes***	***Yes***	*No*
Tools (…)	*Google Earth*	*Web libraries*	*QGIS, ArcGIS*	*QGIS, SpatiaLite*	*QGIS, ArcGIS*	*Big data frameworks*	*ArcGIS, Esri Software*	*CloudCompare*
R	*sf::st_read()*	*sf::st_read()*	*sf::st_read()*	*RSQLite::dbCon()*	*sf::st_read()*	*arrow::read_pq()*	*arcgisbinding::arc.open()*	*lidR::readLAS()*
Python	*gp.read_file()*	*gp.read_file()*	*gp.read_file()*	*sqlite3.connect()*	*gp.read_file()*	*pandas.read_parquet()*	*arcpy*	*laspy.read()*

Spatial Databases

- **PostgreSQL with PostGIS** (Open-source)
  Good for large spatial data with advanced analytics and multi-user capability
- **SpatiaLite** (Open-source)
  Lightweight, portable, ideal for single-user or smaller datasets.
- **MongoDB** (Open-source)
  Best for web-based geospatial applications where scalability, flexibility, and schema-free design are critical
- **DuckDB** (Open-source)
  Ideal for spatial analytics in data science workflows, offering in-memory performance and SQL-based querying.
- **Oracle Spatial**
  Enterprise-level solution with strong topology and network modeling capabilities.
- **SQL Server with Spatial Extension**
  Best for Microsoft-based enterprise environments.
- **Esri Enterprise Geodatabase**:
  Designed for use within the Esri ecosystem for advanced geospatial workflows

`Cloud Optimized` formats

- Designed to be efficient in cloud environments
- Key Features:
  - **Efficient Access**: range-based HTTP requests, reduces bandwidth usage
  - **Self-Describing**: metadata is embedded within the file
  - **Compression**: efficient compression techniques
  - **Interoperability**: can be used by most commercial / open source software
  - **Read-Only**: file is rewritten rather than updated

Format	Type	Optimized For	Use Cases
Cloud Optimized Geotiff (COG)	*Raster*	*Large geospatial rasters*	*Satellite imagery, DEMs*
GeoParquet	*Vector*	*Large-scale vector data*	*Distributed analytics, big data workflows*
Cloud Optimized Point Clouds (COPC)	*Point Cloud*	*LiDAR/3D data in the cloud*	*LiDAR data, terrain modeling, 3D mapping*

Data Sources

Group Discussion

Typical geospatial workflow

1. Data Collection

- Remote Sensing
- Published Datasets
- Field Data

2. Data Storage

- Vector
- Raster
- Databases

3. Data Preprocessing

- Cleaning
- Reprojection
- Transformations

3. Data Analysis

- Spatial Analysis
- Raster Analysis
- Time Series
- Classification
- Statistics

4. Data Visualization

- Static or interactive maps
- Graphs

5. Data Sharing

- Research publication
- Publish streaming services
- Data hosting
- Interactive website

Geospatial Software

- https://github.com/sacridini/Awesome-Geospatial

Examples of advanced geospatial

- https://mghydro.com/watersheds/
- https://sites.research.google/floods/l/0/0/3
- https://firesmoke.ca/forecasts/current/
- https://livingatlas.arcgis.com/landcoverexplorer/
- https://portal.opentopography.org/noaaDataset?noaaID=1451
- https://caltopo.com/map.html

Homework

- **Register**
  - Google Earth Engine (free account) https://earthengine.google.com/ (help)
- **Next Tuesday**
  - Discuss this podcast https://mapscaping.com/podcast/ai-autocomplete-for-qgis/
- **On Thursday**
  - Discuss this paper https://www.mdpi.com/2220-9964/9/2/90