Mine and Flow and Nothing Grow

Spatial Economics Project Pitch

Max Heinze (h11742049@wu.ac.at) Gustav Pirich (h11910449@wu.ac.at)

May 28, 2024

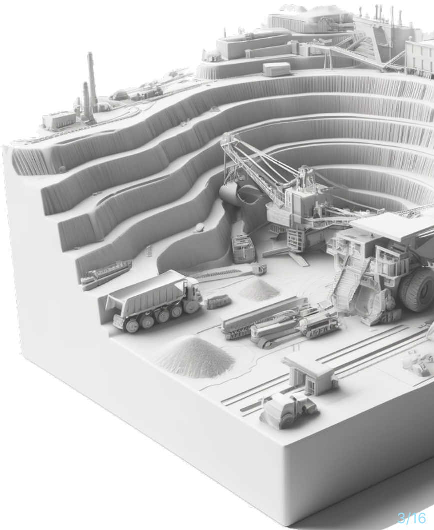
Download Slides

What Do We Want to Find Out?

Data

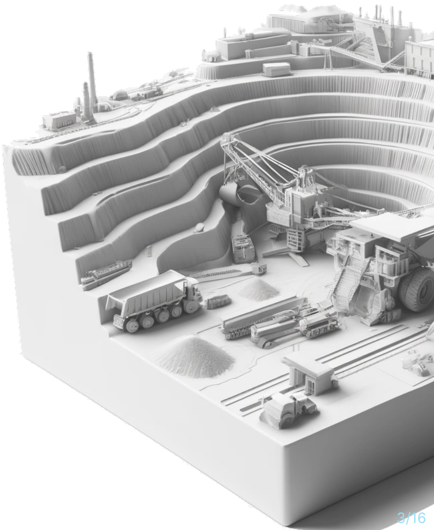
Empirical Framework

Metals Be Mine, Problems Be Thine



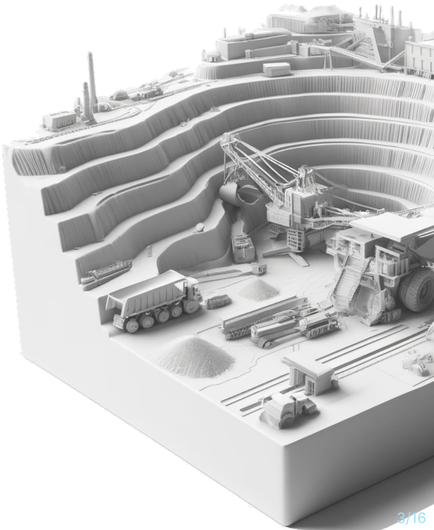
Metals Be Mine, Problems Be Thine

- The African **resource extraction** industry is thriving. Mining activity yields **economic benefits**.



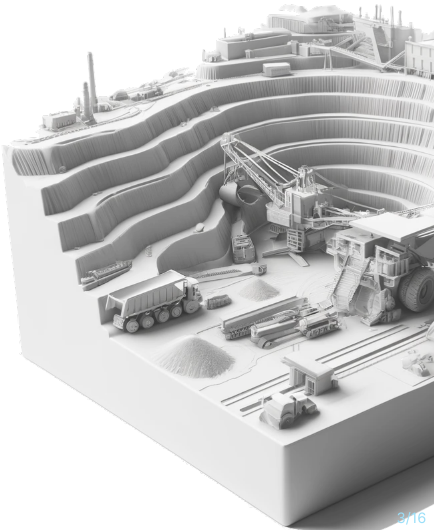
Metals Be Mine, Problems Be Thine

- The African **resource extraction** industry is thriving. Mining activity yields **economic benefits**.
- Mines also generate **negative externalities**, ranging from corruption (Berman et al., 2017) to air-pollution-induced declines in **agricultural productivity** (Aragón & Rud, 2015).

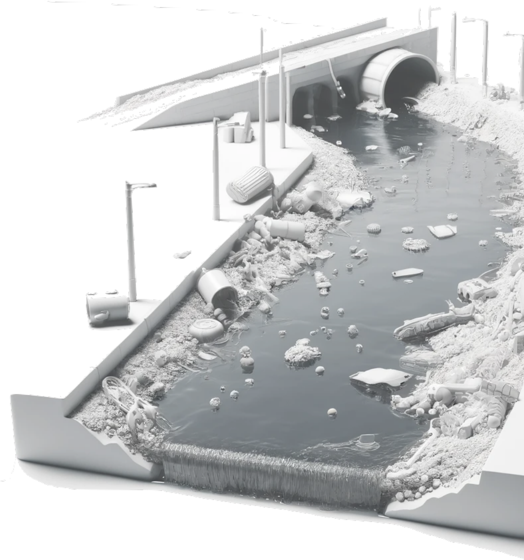


Metals Be Mine, Problems Be Thine

- The African **resource extraction** industry is thriving. Mining activity yields **economic benefits**.
- Mines also generate **negative externalities**, ranging from corruption (Berman et al., 2017) to air-pollution-induced declines in **agricultural productivity** (Aragón & Rud, 2015).
- We want to take a look on how **water pollution** affects **crop yields**.

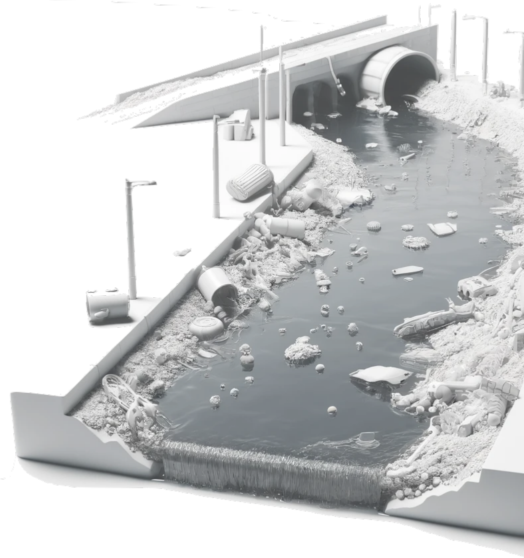


How Does Pollution Travel?



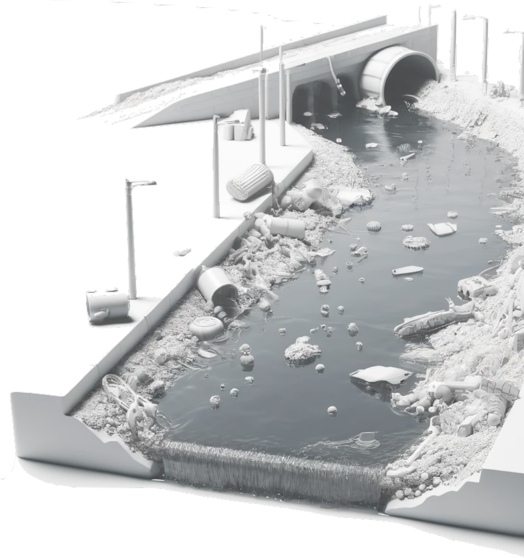
How Does Pollution Travel?

- Mines pollute both the air in their surroundings and adjacent **water flows**.



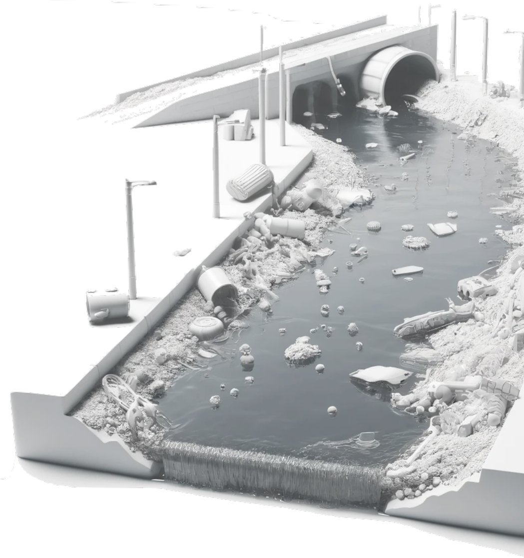
How Does Pollution Travel?

- Mines pollute both the air in their surroundings and adjacent **water flows**.
- The pollutants enter the river at the mining site and **travel with the water**.



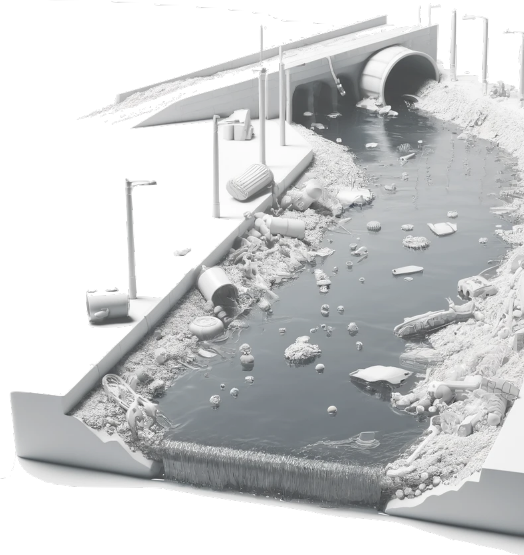
How Does Pollution Travel?

- Mines pollute both the air in their surroundings and adjacent **water flows**.
- The pollutants enter the river at the mining site and **travel with the water**.
- What is the reason that you should not swim in the Danube *east* of Belgrade, the largest European city without wastewater treatment?



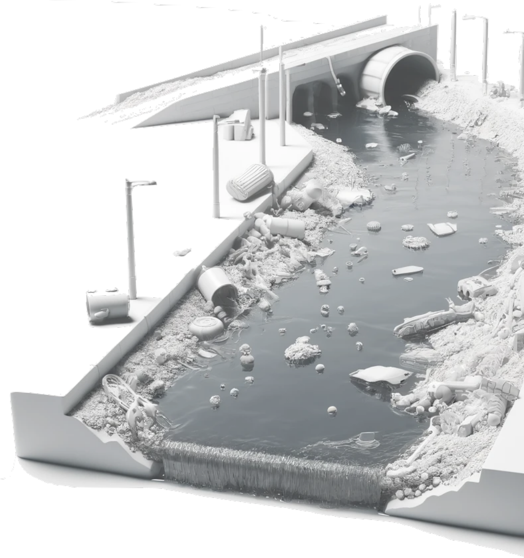
How Does Pollution Travel?

- Mines pollute both the air in their surroundings and adjacent **water flows**.
- The pollutants enter the river at the mining site and **travel with the water**.
- What is the reason that you should not swim in the Danube *east* of Belgrade, the largest European city without wastewater treatment?
- Areas **downstream** are affected by pollution, but areas **upstream** should not!



How Does Pollution Travel?

- Mines pollute both the air in their surroundings and adjacent **water flows**.
- The pollutants enter the river at the mining site and **travel with the water**.
- What is the reason that you should not swim in the Danube *east* of Belgrade, the largest European city without wastewater treatment?
- Areas **downstream** are affected by pollution, but areas **upstream** should not!
- This gives us **treated** and **control** areas.

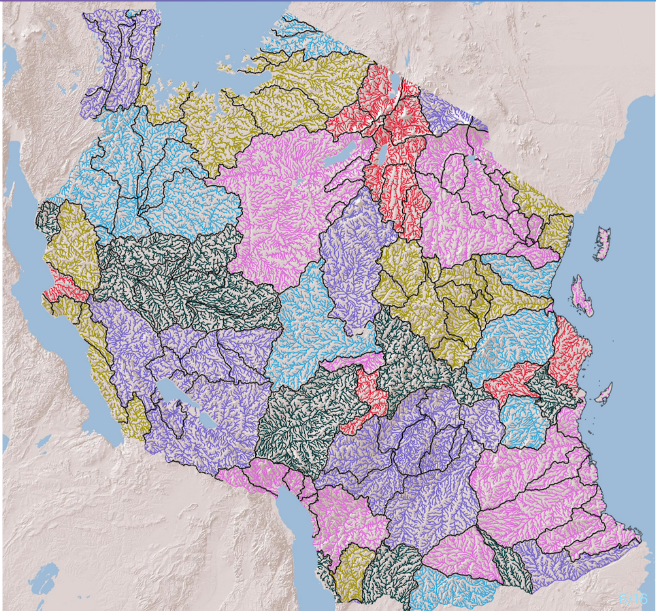


What Do We Want to Find Out?

Data

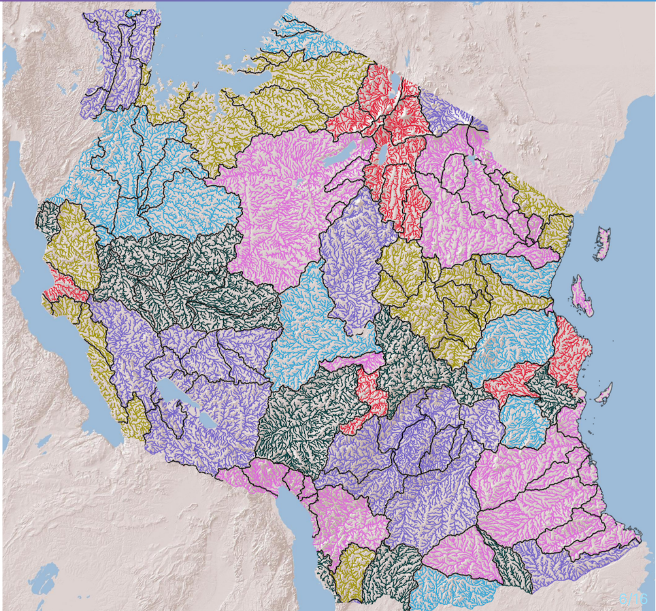
Empirical Framework

Highs, Lows, and Water Flows



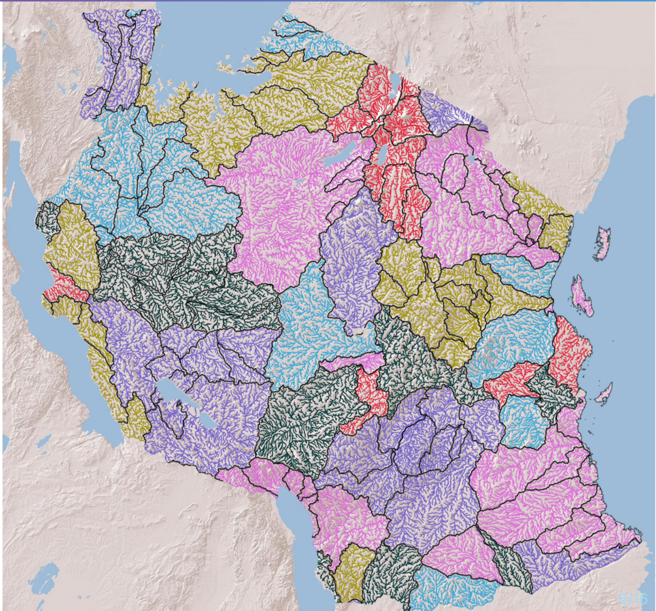
Highs, Lows, and Water Flows

- The **HydroBASINS** dataset from Lehner and Grill (2013) gives us information on where the water drains into in a given region.



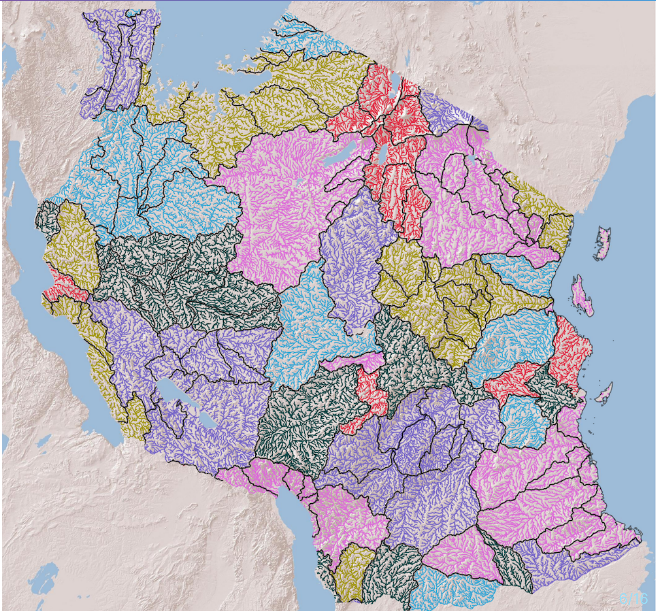
Highs, Lows, and Water Flows

- The **HydroBASINS** dataset from Lehner and Grill (2013) gives us information on where the water drains into in a given region.
- How do they know where water from any point on earth flows to?



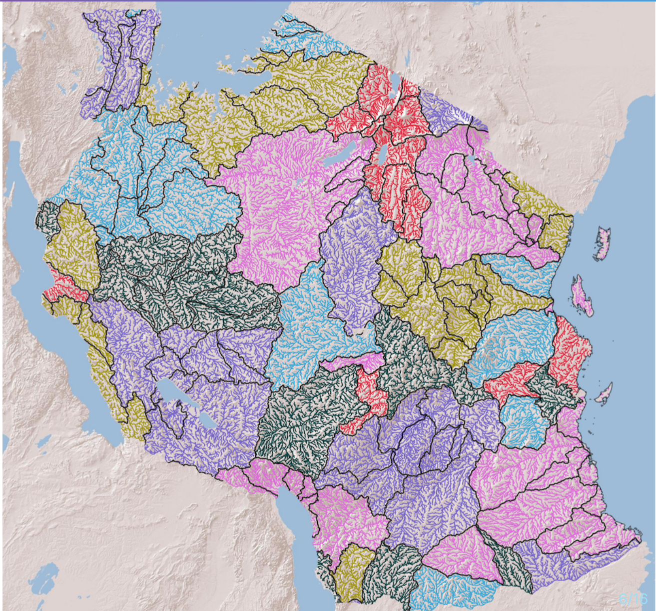
Highs, Lows, and Water Flows

- The **HydroBASINS** dataset from Lehner and Grill (2013) gives us information on where the water drains into in a given region.
- How do they know where water from any point on earth flows to?
- Water flows are predicted from **elevation data**, which stems from the Shuttle Radar Topography Mission.

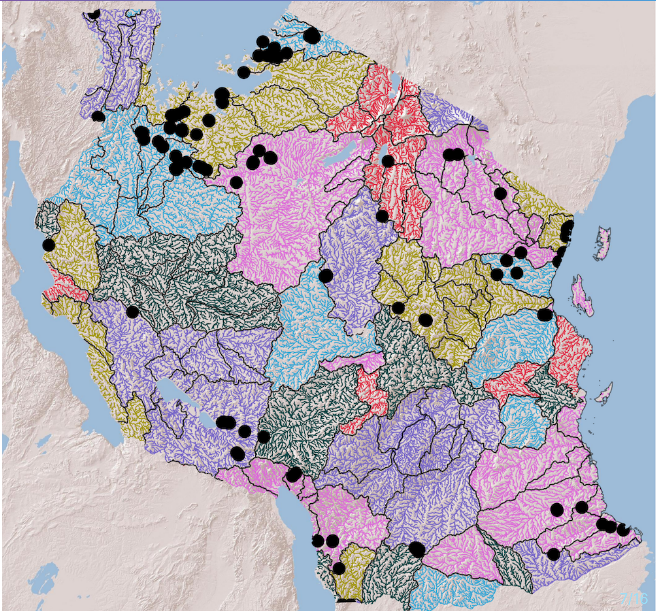


Highs, Lows, and Water Flows

- The **HydroBASINS** dataset from Lehner and Grill (2013) gives us information on where the water drains into in a given region.
- How do they know where water from any point on earth flows to?
- Water flows are predicted from **elevation data**, which stems from the Shuttle Radar Topography Mission.
- This map shows two different levels of **basins** along with **water flows** in them.

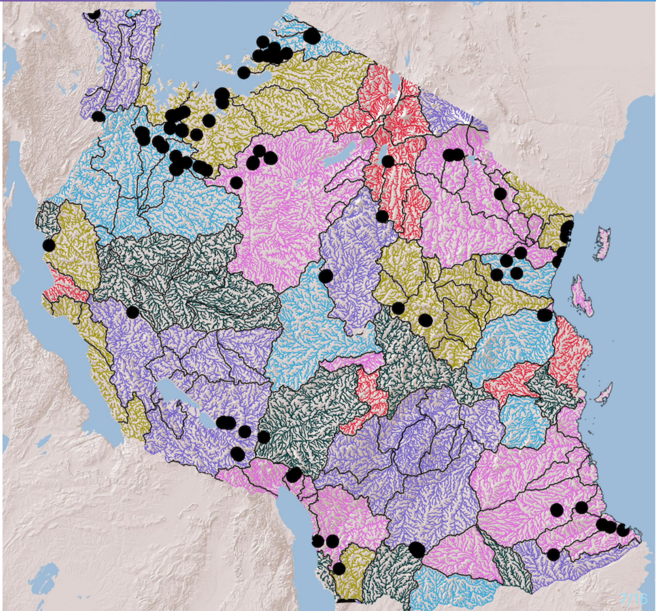


Excuse Me, Do You Know Where the Next Mine Is?



Excuse Me, Do You Know Where the Next Mine Is?

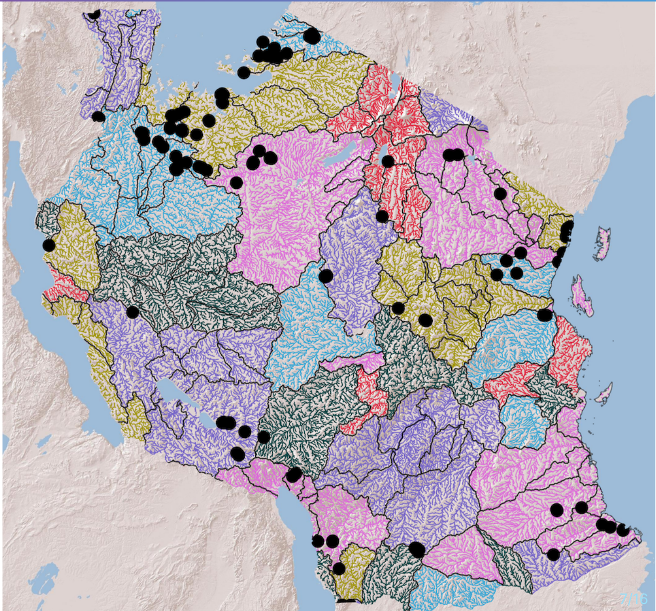
- We can now overlay **centroids of mines**, taken e.g. from Maus et al. (2022) or Kuschnig et al. (2024), over the river basin map.[1]



[1]This map uses Kuschnig et al. (2024).

Excuse Me, Do You Know Where the Next Mine Is?

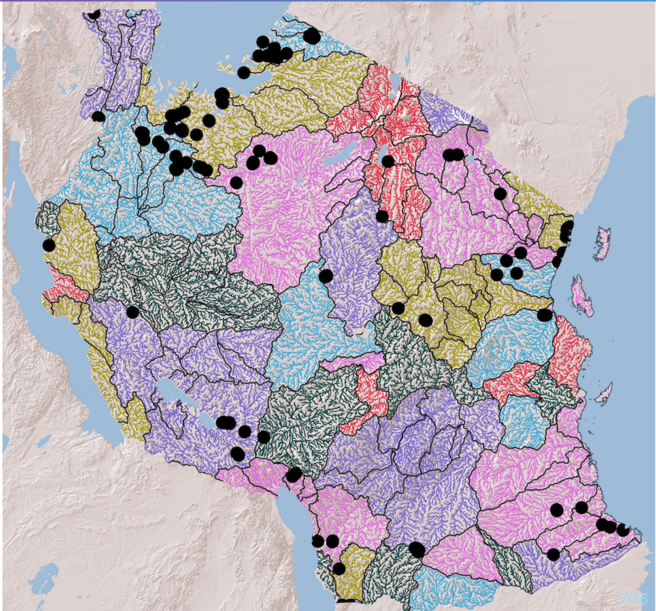
- We can now overlay **centroids of mines**, taken e.g. from Maus et al. (2022) or Kuschnig et al. (2024), over the river basin map.[1]
- How can we relate the basin data with mine locations?



[1]This map uses Kuschnig et al. (2024).

Excuse Me, Do You Know Where the Next Mine Is?

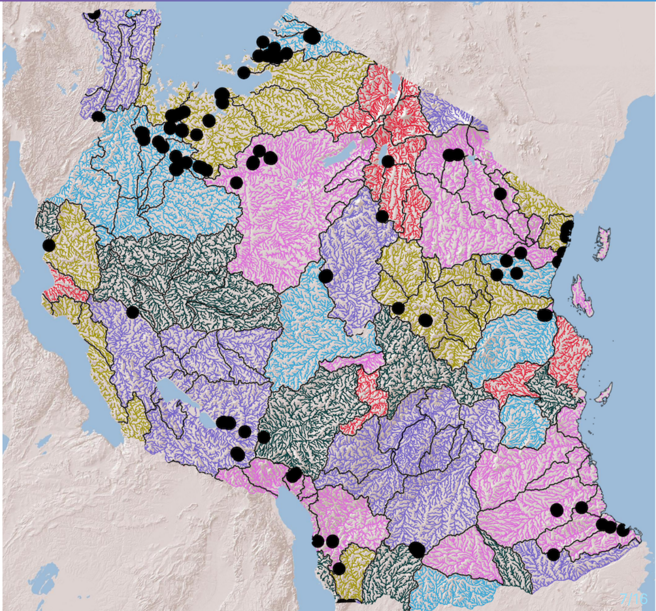
- We can now overlay **centroids of mines**, taken e.g. from Maus et al. (2022) or Kuschnig et al. (2024), over the river basin map.[1]
- How can we relate the basin data with mine locations?
- At the **lowest level**, basins are very granular.



[1]This map uses Kuschnig et al. (2024).

Excuse Me, Do You Know Where the Next Mine Is?

- We can now overlay **centroids of mines**, taken e.g. from Maus et al. (2022) or Kuschnig et al. (2024), over the river basin map.[1]
- How can we relate the basin data with mine locations?
- At the **lowest level**, basins are very granular.
- Basins are labeled according to a **coding system** which allows identification of **up-** and **downstream** basins.



[1]This map uses Kuschnig et al. (2024).

Crop Yields from Space

Crop Yields from Space

- To proxy **agricultural productivity** we can use:

Crop Yields from Space

- To proxy **agricultural productivity** we can use:
    - Annual Maximum Cropland **NDVI/EVI** during growing season (Didan, 2015)

Crop Yields from Space

- To proxy **agricultural productivity** we can use:
  - Annual Maximum Cropland **NDVI/EVI** during growing season (Didan, 2015)
  - Copernicus Crop Yields (Copernicus Climate Change Service, 2020)

Crop Yields from Space

- To proxy **agricultural productivity** we can use:
  - Annual Maximum Cropland **NDVI/EVI** during growing season (Didan, 2015)
  - Copernicus Crop Yields (Copernicus Climate Change Service, 2020)
- We aggregate the proxy for agricultural productivity at the basin level

What Do We Want to Find Out?

Data

Empirical Framework

The Baseline Regression

$$Y_{mdt} = \beta_1 NearMine_{mdt} + \beta_2 Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{1}$$

The Baseline Regression

$$Y_{mdt} = \beta_1 NearMine_{mdt} + \beta_2 Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \quad (1)$$

- *m* is the basin, *d* the district, and *t* the year.

The Baseline Regression

$$Y_{mdt} = \beta_1 NearMine_{mdt} + \beta_2 Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{1}$$

- m is the basin, d the district, and t the year.
- Y_{mdt} measures agricultural productivity.

The Baseline Regression

$$Y_{mdt} = \beta_1 NearMine_{mdt} + \beta_2 Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \quad (1)$$

- m is the basin, d the district, and t the year.
- Y_{mdt} measures agricultural productivity.
- $NearMine_{mdt}$ is an indicator if a basin is in proximity to a mine.

The Baseline Regression

$$Y_{mdt} = \beta_1 NearMine_{mdt} + \beta_2 Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \quad (1)$$

- m is the basin, d the district, and t the year.
- Y_{mdt} measures agricultural productivity.
- $NearMine_{mdt}$ is an indicator if a basin is in proximity to a mine.
- $Downstream_{mdt}$ indicates that the basin is downstream of a mine.

The Baseline Regression

$$Y_{mdt} = \beta_1 NearMine_{mdt} + \beta_2 Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \quad (1)$$

- m is the basin, d the district, and t the year.
- Y_{mdt} measures agricultural productivity.
- $NearMine_{mdt}$ is an indicator if a basin is in proximity to a mine.
- $Downstream_{mdt}$ indicates that the basin is downstream of a mine.
- X_{mdt} are a set of of geographic and socioeconomic controls.

The Baseline Regression

$$Y_{mdt} = \beta_1 NearMine_{mdt} + \beta_2 Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \quad (1)$$

- m is the basin, d the district, and t the year.
- Y_{mdt} measures agricultural productivity.
- $NearMine_{mdt}$ is an indicator if a basin is in proximity to a mine.
- $Downstream_{mdt}$ indicates that the basin is downstream of a mine.
- X_{mdt} are a set of of geographic and socioeconomic controls.
- η_d are district fixed effects.

A (Spatial) Regression Discontinuity Design

$$Y_{mdt} = f(distance_{mdt}) + Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{2}$$

$$Y_{mdt} = f(distance_{mdt}) + f(distance_{mdt}) \times Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{3}$$

A (Spatial) Regression Discontinuity Design

$$Y_{mdt} = f(distance_{mdt}) + Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{2}$$

$$Y_{mdt} = f(distance_{mdt}) + f(distance_{mdt}) \times Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{3}$$

- m is the basin, d the district, and t the year.

A (Spatial) Regression Discontinuity Design

$$Y_{mdt} = f(distance_{mdt}) + Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{2}$$

$$Y_{mdt} = f(distance_{mdt}) + f(distance_{mdt}) \times Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{3}$$

- m is the basin, d the district, and t the year.
- Y_{mdt} measures agricultural productivity.

A (Spatial) Regression Discontinuity Design

$$Y_{mdt} = f(distance_{mdt}) + Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \quad (2)$$

$$Y_{mdt} = f(distance_{mdt}) + f(distance_{mdt}) \times Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \quad (3)$$

- m is the basin, d the district, and t the year.
- Y_{mdt} measures agricultural productivity.
- $f(distance_{mdt})$ measures the distance of the basin's centroid to the nearest mine at the river.

A (Spatial) Regression Discontinuity Design

$$Y_{mdt} = f(distance_{mdt}) + Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{2}$$

$$Y_{mdt} = f(distance_{mdt}) + f(distance_{mdt}) \times Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{3}$$

- m is the basin, d the district, and t the year.
- Y_{mdt} measures agricultural productivity.
- $f(distance_{mdt})$ measures the distance of the basin's centroid to the nearest mine at the river.
- $Downstream_{mdt}$ indicates that the basin is downstream of a mine.

A (Spatial) Regression Discontinuity Design

$$Y_{mdt} = f(distance_{mdt}) + Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{2}$$

$$Y_{mdt} = f(distance_{mdt}) + f(distance_{mdt}) \times Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{3}$$

- m is the basin, d the district, and t the year.
- Y_{mdt} measures agricultural productivity.
- $f(distance_{mdt})$ measures the distance of the basin's centroid to the nearest mine at the river.
- $Downstream_{mdt}$ indicates that the basin is downstream of a mine.
- X_{mdt} are a set of of geographic and socioeconomic controls.

A (Spatial) Regression Discontinuity Design

$$Y_{mdt} = f(distance_{mdt}) + Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{2}$$

$$Y_{mdt} = f(distance_{mdt}) + f(distance_{mdt}) \times Downstream_{mdt} + X_{mdt} + \eta_d + \varepsilon_{mdt} \tag{3}$$

- m is the basin, d the district, and t the year.
- Y_{mdt} measures agricultural productivity.
- $f(distance_{mdt})$ measures the distance of the basin's centroid to the nearest mine at the river.
- $Downstream_{mdt}$ indicates that the basin is downstream of a mine.
- X_{mdt} are a set of of geographic and socioeconomic controls.
- η_d are district fixed effects.

Identification

Identification

Following Cattaneo et al. (2020) and Singh (2022), we can then look at the jump in the conditional expectation of the outcome Y_m given the covariate (distance) X_m at the threshold *Mine*

$$\lim_{x \downarrow Mine} \mathrm{E}(Y_m \mid X_m = x) - \lim_{x \uparrow Mine} \mathrm{E}(Y_m \mid X_m = x)$$

Identification

Following Cattaneo et al. (2020) and Singh (2022), we can then look at the jump in the conditional expectation of the outcome Y_m given the covariate (distance) X_m at the threshold *Mine*

$$\lim_{x \downarrow Mine} \mathrm{E}(Y_m \mid X_m = x) - \lim_{x \uparrow Mine} \mathrm{E}(Y_m \mid X_m = x)$$

which can be interpreted as the average causal effect of treatment at the discontinuity point

$$\tau_{SRD} = \mathrm{E}\left[Y_m(1) - Y_m(0) \middle| X_i = Mine \right]$$

Extensions

Extensions

- We can **match** units based on rivers to compare units at the same river.

Extensions

- We can **match** units based on rivers to compare units at the same river.
- We can expand our **notion of distance** by replacing $f(distance_{mdt}$ with $f(geolocation_{mdt})$ (Dell, 2010).

Extensions

- We can **match** units based on rivers to compare units at the same river.
- We can expand our **notion of distance** by replacing $f(distance_{mdt}$ with $f(geolocation_{mdt})$ (Dell, 2010).
- We might need to control for the presence of **cities** and **factories** (sourced from Open Street Map?).

Extensions

- We can **match** units based on rivers to compare units at the same river.
- We can expand our **notion of distance** by replacing $f(distance_{mdt}$ with $f(geolocation_{mdt})$ (Dell, 2010).
- We might need to control for the presence of **cities** and **factories** (sourced from Open Street Map?).
- We could use **survey data**.

Thank You!

References I

Aragón, F. M., & Rud, J. P. (2015). Polluting Industries and Agricultural Productivity: Evidence from Mining in Ghana. *The Economic Journal*, *126*(597), 1980–2011. https://doi.org/10.1111/ecoj.12244

Berman, N., Couttenier, M., Rohner, D., & Thoenig, M. (2017). This Mine is Mine! How Minerals Fuel Conflicts in Africa. *American Economic Review*, *107*(6), 1564–1610. https://doi.org/10.1257/aer.20150774

Cattaneo, M. D., Titiunik, R., & Vazquez-Bare, G. (2020). The Regression Discontinuity Design. In *The SAGE Handbook of Research Methods in Political Science and International Relations* (pp. 835–857). SAGE Publications Ltd. https://doi.org/10.4135/9781526486387.n47

Copernicus Climate Change Service. (2020). Crop productivity and evapotranspiration indicators from 2000 to present derived from satellite observations. https://doi.org/10.24381/CDS.B2F6F9F6

Dell, M. (2010). *Econometrica*, *78*(6), 1863–1903. https://doi.org/10.3982/ecta8121

Didan, K. (2015). MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. https://doi.org/10.5067/MODIS/MOD13Q1.006

Kuschnig, N., Sepin, P., & Vashold, L. (2024). *Global mapping and delineation of mining areas*. [Unpublished Manuscript].

Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrological Processes*, *27*(15), 2171–2186. https://doi.org/10.1002/hyp.9740

References II

Maus, V., da Silva, D. M., Gutschlhofer, J., da Rosa, R., Giljum, S., Gass, S. L. B., Luckeneder, S., Lieber, M., & McCallum, I. (2022). *Global-scale mining polygons (Version 2)* (Dataset). PANGAEA. https://doi.org/10.1594/PANGAEA.942325

Singh, A. (2022). RDD.