Thesis Proposal

Eye in the Sky
Exploring the Socio-Economic and Ecological Impact of Mining and Heavy Industries Using Satellite Imagery

Sayan Dasgupta

Department of Economics
Presidency University

November 2024

1 Introduction

The extraction of natural resources is fundamental to sustainable global economic and industrial development, yet its environmental and social costs are increasing at an alarming rate. Africa is a critical player in the energy transition and urban development due to its abundant reserves of energy transition minerals (ETM). ETMs are naturally occurring substances that are ideal for use in the renewable energy sector. The Democratic Republic of Congo (DRC) holds 54% of the world's cobalt reserves
(World Bank 2021), making it a critical component in the production of lithium-ion batteries essential for renewable energy technologies.In 2019, Guinea was the world's third-ranked producer of bauxite, accounting for 19% of world output, vital for aluminum production (U.S. Geological Survey 2019), while South Africa, Gabon and Ghana collectively account for over 60 percent of global manganese production. Zimbabwe, alongside the Democratic Republic of Congo and Mali, hold substantial but yet-to-be-explored lithium deposits. Other countries with significant critical mineral reserves include Guinea, South Africa, and Zambia.(IMF 2024,Harnessing Sub-Saharan Africa's Critical Mineral Wealth).

However, the rapid pace of resource extraction raises severe socio-economic and environmental concerns. Artisanal and informal mining dominate in many regions in Africa, providing livelihoods but perpetuating precarious informal economies and labor exploitation. Environmentally, the consequences are dire: deforestation, biodiversity loss, water contamination, and soil degradation are prevalent across mining zones.

Adding to these challenges, sand mining has emerged as a critical issue. With global sand consumption exceeding 50 billion tons annually (World economic forum 2023), primarily driven by construction, unregulated extraction has led to ecological destruction and economic disarray (UN Environment Programme 2022). Sub-Saharan Africa is witnessing rampant illegal sand mining, destabilizing river systems, and accelerating groundwater depletion. Coastal sand extraction in North Africa has increased saltwater intrusion into freshwater aquifers, jeopardizing agriculture and drinking water supplies. Alarmingly, the global market value of illegal sand extraction ranges from \$199.88 billion to \$349.98 billion annually, representing a substantial economic and ecological cost (L.F. Ramadon ,2024).

The unchecked extraction of sand and ETMs is fueling an alarming increase in conflicts and wars in affected regions. As global demand for resources like sand, cobalt, and lithium surges, local communities often find themselves in direct conflict with mining corporations, state actors and organized criminal groups(International Crisis Group). Similarly, violent clashes linked to sand mining in Africa is quite frequent and often results in loss of human life (Aduda et al. 2024).

These resource-driven conflicts disproportionately affect indigenous communities, exacerbating socioeconomic inequalities while undermining peace-building efforts. The environmental degradation caused by mining—such as contaminated waterways, deforestation, and habitat destruction—further amplifies tensions by threatening livelihoods dependent on agriculture, fisheries, and tourism. Urgent interventions are required to mitigate these escalating conflicts and address the underlying drivers of resource-related violence.

Despite the complexity of these challenges, advances in satellite-based monitoring provide opportunities for large-scale environmental and socioeconomic assessments. This study integrates remote sensing technologies with socioeconomic data to examine the dual impact of

Energy Transition Minerals (ETMs) and sand extraction, investigating their role in driving economic development while also assessing their environmental and socioeconomic consequences. Through this approach, actionable recommendations can be proposed to balance resource demands with sustainable practices, ensuring equitable socioeconomic growth while protecting ecological integrity.

2 Literature Review and Research Gaps

2.1 Socioeconomic and Environmental Impacts of Energy Transition Minerals (ETMs) as seen through a mix of satellite imagery and socioeconomic data

(Madasa et al.,2021),(Orimoloye et al.,2020) demonstrated that remote sensing technologies can effectively map land-use changes in mining areas, identifying significant spatial transformations, rapid landscape modifications, and complex environmental dynamics.
(Shackleton,2020) documented substantial land loss and livelihood disruptions in Limpopo, South Africa. (Fotheringham et al.,2002) revealed emerging land-grabbing practices, complex social stratification, and economic marginalization of local communities. (Breiman,2001) identified increased environmental vulnerability and disruption of traditional social structures.
(Hertel et al.,1997) highlighted tensions between mining for economic development and significant environmental and cultural trade-offs. (Yiran et al.,2012) emphasized the importance of integrating remote sensing with local knowledge, developing context-specific degradation assessment methods, and recognizing local perspectives in environmental research.

The environmental consequences of ETM mining have been extensively studied, with notable works such as (Werner et. al. 2019), (Rudke et. al. 2020) and (Paul et. al. 2006). However, the socioeconomic and environemntal impacts of proximity to ETM mines, particularly in countries such as the Democratic Republic of Congo (DRC), Zimbabwe, Zambia, Mozambique, Tanzania and Gabon, remain under-explored.

Research combining satellite imagery with socioeconomic data to study these impacts is even scarcer. Between 1996 and 2021, only five papers addressing this intersection in Africa have been published. These studies predominantly focus on localized regions, typically involving one or two mines. A recent systematic review (Ang et al., 2023) highlights the limited yet growing application of GIS and remote sensing techniques in assessing the socioeconomic impacts of mining activities, emphasizing the potential of these methodologies for broader, more comprehensive studies.

2.2 Illegal Sand Mining and Its Socioeconomic Impacts in West Africa

2.2.1 Remote Sensing and Mapping Techniques

Recent advancements in Remote Sensing (RS) and Geographic Information Systems (GIS) have significantly improved the capability to detect and map illegal sand mining sites.

(Duan et al ,2019) developed a comprehensive methodology using multi-spectral satellite imagery and machine learning algorithms to identify sand mining hotspots in China's fourth largest freshwater lake, demonstrating the potential of advanced geospatial technologies in monitoring illegal extractive activities. (Gallwey et al. ,2020) introduced an innovative approach combining Sentinel-2 satellite data with deep learning techniques, specifically multi-spectral U-net convolutional neural networks (CNNs), to automatically detect and classify artisanal sand mining sites in Ghana. Their research highlighted the effectiveness of integrating high-resolution satellite imagery with artificial intelligence for large-scale environmental monitoring.

2.2.2 Socioeconomic Impacts

The socioeconomic dimensions of illegal sand mining remain complex and multifaceted. (Johnbull et al. ,2017) has worked on socioeconomic consequences of sandmining along the Victory river in Nigeria. This work has revealed : Illegal sand mining provides critical employment opportunities in regions with limited economic alternatives. The sector supports a broader informal economy, including transportation, small-scale trading, and associated services. Youth unemployment remains a significant driver of participation in illegal sand mining activities. But the uncontrolled extraction of sand along the Victory River in Nigeria, has led to ecosystem destruction, biodiversity loss, and disruption of livelihoods, despite its short-term economic benefits.

A study by Rahman et al. (2023) in coastal regions highlighted the economic paradox: while sand mining provides short-term economic benefits, it simultaneously undermines long-term economic sustainability through environmental degradation.

Sand mining activities in Benin City, Edo State, Nigeria, have led to significant environmental degradation, with satellite imagery and DEM data highlighting the spatial and volumetric impacts of mining sites. Sustainable mining practices and stricter enforcement of Environmental Impact Assessments (EIA) are essential to mitigate these effects (Asikhia et al .,2021).

2.3 Environmental and Socioeconomic Costs of Heavy Industries

Multitemporal Nighttime Light (NTL) data has been increasingly utilized to study the socioeconomic impacts of heavy industries, including regional inequality, energy consumption, and environmental degradation (Xu et al., 2015; Zhou et al., 2015). For instance, NTL data has been leveraged to monitor the effects of industrial expansion on local ecosystems and urbanization (Geldman et al., 2014; Liu et al., 2014). In regions with heavy industries, NTL can reflect changes in electricity consumption patterns and infrastructure development, helping to assess the environmental and developmental consequences of industrial activity (Cao et al., 2013; Shi et al., 2015). Additionally, NTL data has been employed to track the human impacts of large-scale industrial operations, such as resource extraction and steel manufacturing, and their effects on surrounding communities (Liang et al., 2014). While NTL data is useful for understanding industrial dynamics, it is also essential to consider regional variations in light emissions, influenced by factors like industrial density, lighting regulations, and socioeconomic factors (Kyba et al., 2015). Integrating NTL data with other remotely sensed datasets, such as MODIS or Landsat, can provide a more comprehensive view of how industrial activities interact with environmental and social factors (Levin & Duke, 2012; Keola et al., 2015). Site-specific analyses, including fieldwork, can further reveal how local industrial practices influence NTL patterns and contribute to environmental change (M.M. Bennett & L.C. Smith, 2017).

Research Gaps

1. While the environmental consequences of ETM mining are somewhat documented, little research has been conducted to analyze the direct socio-economic effects in specific regions like DRC, Zimbabwe, Zambia, Mozambique, Tanzania, and Gabon. There is an opportunity to bridge this gap by combining satellite imagery with socioeconomic data to explore the broader impacts of ETM extraction on local communities.

2. The application of RS and GIS has enhanced the ability to monitor and analyze the impacts of illegal and legal sand mining activities. However, gaps remain in large scale reporting of illegal sand mining sites in real time and understanding the socio-economic impacts of illegal sand mining activities:

    - **Integrating socio-economic data:** Combining environmental data with socio-economic analyses to better understand the broader consequences of illegal mining.
    - **Groundwater monitoring:** Expanding the use of Gravity Recovery and Climate Experiment (GRACE) data for detailed monitoring of groundwater resources in mining areas, especially in areas where illegal mining is prevalent.
    - **Longitudinal studies:** Long-term studies tracking the cumulative effects of illegal sand mining, particularly in coastal and riverine environments.

3. The environmental and economic consequences of heavy industries like steel and aluminum production in regions such as Gabon and Zambia need further examination. There is a specific gap in quantifying these impacts using satellite imagery, which could offer a unique perspective on the spatial distribution and intensity of the effects on local communities and the environment.

3 Research Objectives

The primary objectives of this research are as follows:

1. To analyze the socioeconomic and environmental impacts of increase in demand of Energy Transition Minerals (ETMs) in regions including the DRC, Zimbabwe, Zambia, Mozambique, Tanzania, and Gabon via satellite imagery. This includes effects on local livelihoods, infrastructure, ecosystem degradation.

2. To develop a framework for monitoring illegal sand mining activities in West and sub Saharan Africa using satellite imagery.

3. To evaluate the socio-economic impacts of sand extraction on West African coastal communities, including its role in informal economies and contributions to coastal erosion.

4. To quantify the socioeconomic and environmental costs of heavy industries in Africa using satellite imagery.

4 Data and Methods

Industrialization, driven by humanity's pursuit of sustainable development, has led to a significant increase in ETM demand. This section outlines the comprehensive methodology and data sources employed to assess the socio-economic and environmental impacts of increase in demand of Energy Transition Minerals (ETMs) in various African regions, monitor illegal sand mining activities, and quantify the costs on local communities and ecosystems. We combine advanced statistical models, remote sensing data, and socio-economic data from several sources to provide a detailed and dynamic understanding of these issues.

4.1 Objective 1: ETM Impact Analysis in African Regions

To begin our exploration of the socio-economic impacts of Energy Transition Minerals (ETMs), we employ a **spatial-temporal impact assessment** approach. This includes the use of a **Difference-in-Differences (DiD)** regression model, which helps us capture the dynamic changes before and after the extraction of ETMs or before and after expansion in extraction of ETM from mines in specific regions. Data for this model is sourced from reputable institutions such as the World Bank, IMF, and national statistical agencies, along with regional economic databases like the African Development Bank (AfDB).

Next, we incorporate **Geographically Weighted Regression (GWR)** to account for spatial heterogeneity in the data. This approach allows us to understand how the impacts of ETM extraction vary across different geographical locations. We leverage remote sensing data from sources like Landsat, MODIS, and other regional GIS databases, in addition to socio-economic datasets from national surveys such as the Demographic and Health survey, to capture regional variations in the effects of ETM extraction.

For the economic impacts, we turn to **Computable General Equilibrium (CGE) modeling** to quantify the broader economic effects of ETM extraction. Using national input-output tables and country-level economic data (e.g., from UN Comtrade and national economic surveys), we construct a model that simulates how ETM extraction interacts with various economic sectors. Additionally, we apply a **Structural Transformation Economic Model** to explore the shifts in industry structure and the role of investment in mining and infrastructure development.

Satellite imagery plays a crucial role in understanding environmental changes due to ETM extraction. We use **Random Forest Classification** (Billah et al. ,2023) to analyze land-use changes, supported by imagery from Landsat, Sentinel-2, and MODIS. Time-series analysis of the **Normalized Difference Vegetation Index (NDVI)** allows us to monitor vegetation health over time, providing insights into how mining activities may influence local ecosystems. Finally, we calculate a **Landscape Fragmentation Index** to assess the spatial structure of the land and the effects of fragmentation caused by mining activities.

4.2 Objective 2: Illegal Sand Mining Monitoring Framework

To detect and monitor illegal sand mining, we utilize a variety of methods, beginning with **machine learning detection models**. Specifically, we employ a **Convolutional Neural Network (CNN)** architecture to automatically identify mining activities in satellite images. The data for training the CNN comes from high-resolution satellite imagery, including sources like Sentinel-1, Sentinel-2, and WorldView, as well as labeled mining datasets and land-use databases.

In addition, we use the **U-Net Semantic Segmentation Model** to segment and detect mining features in remote sensing imagery. This model is trained on labeled high-resolution satellite data, such as from WorldView and GeoIQ. Another technique is **Object Detection using YOLO (You Only Look Once)**, a real-time object detection method, which helps to identify mining-related features in satellite images.

Furthermore, we employ **multi-temporal change detection algorithms** to analyze how mining activities evolve over time. Using time-series satellite data, we can detect land-use changes associated with sand extraction. **Spectral Mixture Analysis (SMA)** is also used to separate the different land-cover types, aiding in the detection of sand mining activities. Additionally, we apply the **Normalized Difference Water Index (NDWI)** transformation to identify changes in water bodies that may be associated with mining activities.

4.3 Objective 3: Socio-Economic Impacts of Sand Extraction

In analyzing the socio-economic impacts of illegal sand extraction, we utilize **econometric modeling** techniques. A key tool is the **Panel Vector Autoregression (Panel VAR)** model, which enables us to examine the dynamic interactions between regional socio-economic indicators, such as income and employment rates, and environmental data derived from satellite imagery.This allows us to understand how illegal sand mining activities influence and are influenced by these indicators over time. We also use the **Spatial Durbin Model** to study the effects of sand extraction on coastal communities.

Another important tool is **Propensity Score Matching (PSM)**, which helps us estimate causal effects by comparing regions affected by illegal sand mining with similar, unaffected regions. This method helps ensure that the estimated impacts are not biased by confounding variables. The data for these models comes from socio-economic surveys, environmental datasets, and industrial impact reports.

4.4 Objective 4: Quantifying Industrial Environmental Costs

To evaluate the environmental costs of industrialization, we take a multi-dimensional approach. We begin with **environmental cost estimation models**, such as the **Integrated Environmental-Economic Accounting Framework**, which allows us to estimate the economic value of environmental losses. We also apply a **Spatial Hedonic Pricing Model** to measure how proximity to industrial activities affects property values, using real estate data, air and water quality indicators, and proximity to industrial sites.

Finally, satellite-based environmental metrics play a key role in assessing the environmental impacts of industrial activities. For example, **Emission Proxy Estimation using Spectral Indices** allows us to estimate air pollution levels based on satellite imagery, while **land transformation and degradation tracking** enables us to monitor changes in land cover and vegetation associated with industrial activities. Additionally, we apply **multi-spectral industrial footprint analysis** to assess the spatial extent of industrial impacts using high-resolution satellite imagery and industrial activity data.

References

[World Bank 2021] cobalt in the democratic republic of congo

[U.S. Geological Survey 2019] USGS (2019). The Mineral Industry of Guinea

[IMF 2024,Harnessing Sub-Saharan Africa's Critical Mineral Wealth] IMF (2024).

[UN Environment Programme 2022] United Nations Environment Programme (UNEP) (2022). Sand mining and its ecological and economic impacts.

[World economic forum 2023] World Economic Forum. "Global sand mining is causing environmental harm — and a shortage of this crucial resource." *World Economic Forum*, September 2023.

[L.F. Ramadon ,2024] L. F. Ramadon, "The global estimated value of illegal sand extraction," Guest post, Oct. 29, 2024. Federal Police Specialist in Brazil. Master's degree in Management and Regulation of Water Resources (PROFÁGUA/UERJ) and Post-graduate degree in Environmental Law. [Online]. Available: `http://lattes.cnpq.br/5094490404365005`

[International Crisis Group] International Crisis Group. `https://www.crisisgroup.org`.

[Aduda et al. 2024] L. Aduda and L. Bolf, "The conflict potential of sand: Illegal sand mining on the African continent," *Journal of Sand Studies*, vol. 2, no. 4, 2024. [Online]. Available: `https://doi.org/10.1177/27538796241230583`

[timesofindia2022sandconflict] Times of India. India's Sand Mining Mafias: Violence, Corruption, and Deaths on the Rise. Published: September 2022. Available at: `https://timesofindia.indiatimes.com`.

[Werner et. al. 2019] T. T. Werner, A. Bebbington, and G. Gregory, "Assessing impacts of mining: Recent contributions from GIS and remote sensing," *The Extractive Industries and Society*, vol. 6, no. 3, pp. 993–1012, 2019.

[Rudke et. al. 2020] A. P. Rudke, V. A. S. de Souza, A. M. Dos Santos, A. C. F. Xavier, O. C. Rotunno Filho, and J. A. Martins, "Impact of mining activities on areas of environmental protection in the southwest of the Amazon: A GIS-and remote sensing-based assessment," *Journal of Environmental Management*, vol. 263, p. 110392, 2020.

[Paul et. al. 2006] D. Paull, G. Banks, C. Ballard, and D. Gillieson, "Monitoring the environmental impact of mining in remote locations through remotely sensed data," *Geocarto International*, vol. 21, no. 1, pp. 33–42, 2006.

[Ang et al. ,2023] Michelle Li Ern Ang, John R. Owen, Christopher N. Gibbins, Eléonore Lèbre, Deanna Kemp, Muhamad Risqi U. Saputra, Jo-Anne Everingham, and Alex M. Lechner. "Systematic Review of GIS and Remote Sensing Applications for Assessing the Socioeconomic Impacts of Mining." *The Journal of Environment & Development*, vol. 32, no. 3, 2023, pp. 243–273. DOI: 10.1177/10704965231190126.

[Billah et al. ,2023] Maruf Billah, A.K.M. Saiful Islam, Wasif Bin Mamoon, and Mohammad Rezaur Rahman, "Random forest classifications for landuse mapping to assess rapid flood damage using Sentinel-1 and Sentinel-2 data," *Remote Sensing Applications: Society and Environment*, vol. 30, pp. 100947, 2023. doi: https://doi.org/10.1016/j.rsase.2023.10094710.1016/j.rsase.2023.100947.

[Ferring et al.,2019] Ferring, D., & Hausermann, H. "The political ecology of landscape change, malaria, and cumulative vulnerability in Central Ghana's gold mining country." *Annals of the American Association of Geographers*, vol. 109, no. 4, 2019, pp. 1074–1091. DOI: https://doi.org/10.1080/24694452.2018.1552086.

[Githiria et al.,2020] Githiria, J. M., & Onifade, M. "The impact of mining on sustainable practices and the traditional culture of developing countries." *Journal of Environmental Studies and Sciences*, vol. 10, no. 4, 2020, pp. 394–410. DOI: https://doi.org/10.1007/s13412-020-00624-9.

[Hausermann et al.,2018] Hausermann, H., Ferring, D., Atosona, B., et al. "Land-grabbing, land-use transformation and social differentiation: Deconstructing 'small-scale' in Ghana's recent gold rush." *World Development*, vol. 108, 2018, pp. 103–114. DOI: https://doi.org/10.1016/j.worlddev.2018.03.014.

[Madasa et al.,2021] Madasa, A., Orimoloye, I. R., & Ololade, O. O. "Application of geospatial indices for mapping land cover/use change detection in a mining area." *Journal of African Earth Sciences*, vol. 175, 2021, article 104108. DOI: https://doi.org/10.1016/j.jafrearsci.2020.104108.

[Orimoloye et al.,2020] Orimoloye, I. R., & Ololade, O. O. "Spatial evaluation of land-use dynamics in gold mining area using remote sensing and GIS technology." *International Journal of Environmental Science and Technology*, vol. 17, no. 11, 2020, pp. 4465–4480. DOI: https://doi.org/10.1007/s13762-020-02758-5.

[Shackleton,2020] Shackleton, R. T. "Loss of land and livelihoods from mining operations: A case in the Limpopo province, South Africa." *Land Use Policy*, vol. 99, 2020, article 104825. DOI: https://doi.org/10.1016/j.landusepol.2020.104825.

[Yiran et al.,2012] Yiran, G. A. B., Kusimi, J. M., and Kufogbe, S. K. "A synthesis of remote sensing and local knowledge approaches in land degradation assessment in the Bawku East District, Ghana." *International Journal of Applied Earth Observation and Geoinformation*, vol. 14, no. 1, 2012, pp. 204–213. DOI: `https://doi.org/10.1016/j.jag.2011.09.015`.

[Athey et al.,2006] Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, 74(2), 431-497.

[Fotheringham et al.,2002] Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). Geographically weighted regression: The analysis of spatially varying relationships. *John Wiley & Sons*.

[Hertel et al.,1997] Hertel, T. W. (Ed.). (1997). Global trade analysis: Modeling and applications. *Cambridge University Press*.

[Breiman,2001] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.

[Liu et al,2019] Liu, P., Choo, K.K.R., Wang, L., et al. (2017). SVM or deep learning? A comparative study on remote sensing image classification. *Soft Computing*, 21, 7053–7065. https://doi.org/10.1007/s00500-016-2247-2

[Ronneberger et al,2015] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234-241.

[Abrigo et al ,2016] Abrigo, M. R. M., & Love, I. (2016). Estimation of panel vector autoregression in Stata. *The Stata Journal*, 16(3), 778-804.
‘

[Rosenbaum et al,1983] Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.

[Uno et al,2007] Uno, K., & Bartelmus, P. (2007). Environmental accounting in theory and practice. *International Journal of Environment and Sustainable Development*.

[Duan et al ,2019] Duan, H., Cao, Z., Shen, M., Liu, D., & Xiao, Q. (2019). Detection of illicit sand mining and the associated environmental effects in China's fourth largest freshwater lake using daytime and nighttime satellite images. *Science of the Total Environment*, *647*, 606–618.

[Gallwey et al. ,2020] Gallwey, J., Robiati, C., Coggan, J., Vogt, D., & Eyre, M. (2020). A Sentinel-2 based multispectral convolutional neural network for detecting artisanal small-scale mining in Ghana: Applying deep learning to shallow mining. *Remote Sensing of Environment*, *248*, 111970.

[Johnbull et al. ,2017] Johnbull, S. W., & Brown, I. (2017). Socioeconomic consequences of sand mining along the Victory River in Portharcourt, Nigeria. *Asian Journal of Environment and Ecology*, *3*(2), 1–15.

[Rahman et al.,2023] Rahman, M. A., & Suykens, B. (2023). Sand urbanism in Bangladesh: Transitions of sand extraction and trade in Dhaka-Narayanganj. *The Extractive Industries and Society*, *14*, 101265. https://doi.org/10.1016/j.exis.2023.101265.

[Asikhia et al .,2021] Asikhia, M. O., & Oluku, S. (2021). Geospatial assessment of the impacts of sand mining activities in Benin City, Edo State Nigeria. *Journal of Geography, Environment and Earth Science International*, *25*(1), 46–57.

[Xu et al,2015] Xu, Z., & Zhou, Y. (2015). Regional inequality and economic dynamics from nighttime light data. *Remote Sensing of Environment, 164*, 138-148. https://doi.org/10.1016/j.rse.2015.03.017

[Zhou et al,2015] Zhou, Y., & Xu, Z. (2015). Using nighttime light data to study regional inequality and energy consumption patterns. *Energy Economics, 52*, 100-111. https://doi.org/10.1016/j.eneco.2015.01.014

[Geldman et al,2014] Geldman, J., & Liu, J. (2014). Human impacts on the environment in the context of industrial growth and NTL data. *Journal of Environmental Management, 149*, 163-172. https://doi.org/10.1016/j.jenvman.2014.06.031

[Li et al,2014] Li, F., & Li, Q. (2014). Effects of armed conflict on the built environment and its relationship with NTL data. *Global Environmental Change, 29*, 11-21. https://doi.org/10.1016/j.gloenvcha.2014.07.007

[Venter et al,2016] Venter, O., Geldmann, J., & Gaston, K. J. (2016). Multitemporal NTL as a tool for understanding conservation impacts. *Biological Conservation, 197*, 154-162. https://doi.org/10.1016/j.biocon.2016.03.016

[Gaston et al,2015] Gaston, K. J., Geldmann, J., & Venter, O. (2015). How expanding human footprints impact protected areas: Insights from NTL data. *Conservation Biology, 29*(4), 1015-1024. https://doi.org/10.1111/cobi.12510

[Witmer et al,2011] Witmer, F. D., & O'Loughlin, J. (2011). Using NTL data to examine refugee movements and large-scale fires in the Caucasus. *International Journal of Remote Sensing, 32*(16), 4497-4514. https://doi.org/10.1080/014311

Appendix : Models and Variables

Objective 1: ETM Impact Analysis in African Regions

Difference-in-Differences (DiD) Regression Model

$$Y_{it} = \alpha + \beta_1 D_{it} + \beta_2 T_{it} + \beta_3 (D_{it} \times T_{it}) + X_{it}\gamma + \epsilon_{it}$$

- **Environmental dependent variable:**
  - NDVI (vegetation health).

  - Surface temperature anomalies (e.g., urban heat islands from mining infrastructure).
  - Water body extent or quality (e.g., turbidity in nearby rivers/lakes).
  - Air Quality

- **Socioeconomic Proxies for dependent variable(Indirect from Satellite Imagery):**
  - Night-time light intensity (proxy for economic activity).
  - Land-use changes (urban expansion, deforestation, or infrastructure development).
  - Road and transportation network mapping (High-resolution data from Planet Labs or Sentinel-2 ,historical mapping with Landsat or SAR)

  Treatment Variable (D_{it}):

  - Presence of mining activity can be identified via:
    * Mine locations from geospatial datasets (e.g., USGS or national mining agencies).
    * Changes in surface reflectance in mining areas using Landsat or Sentinel data.

  Time Period (T_{it}):

  - Use satellite imagery from periods before and after mining activities began.
  - Historical Landsat (since the 1970s) or Sentinel data can provide the necessary temporal resolution.

  Control Variables (X_{it}):

  - Environmental covariates derived from satellite data:
    * Rainfall (CHIRPS or IMERG data).
    * Elevation or slope (from SRTM or DEMs).
    * Soil moisture (SMAP or similar).
    * Deforestation rates (from Hansen Global Forest Change dataset).

Example Satellite-Driven DiD Workflow

1. **Identify Mining Areas (Treatment) and Control Areas:**
   - Define polygons for areas impacted by ETM extraction and comparable unaffected regions.
   - Use datasets like mining concession shapefiles or mine location points.

2. **Collect Satellite-Derived Metrics:**
   - Download data for the desired outcome variables (e.g., NDVI, night-time lights) using Google Earth Engine, NASA EarthData, or similar platforms.

3. **Pre-Processing:**
    - Ensure cloud-free imagery for environmental data.
    - Standardize and normalize metrics like NDVI or radiance values.
4. **Temporal Aggregation:**
    - Calculate averages for pre-mining (baseline) and post-mining periods.

Tools for Satellite Data Analysis

- **Google Earth Engine (GEE):** Ideal for large-scale temporal and spatial analysis.
- **QGIS/ArcGIS:** For manual geospatial analysis.
- **Python/R:**
    - Libraries like `rasterio`, `numpy`, and `geopandas` (Python) or `raster` and `terra` (R).

Challenges

- **Economic Metrics:** Satellite proxies (like night-time lights) are indirect measures and may require validation with ground data.
- **Spatial Resolution:** Ensure the resolution of the satellite data matches the scale of the mining impact.
- **Data Availability:** Some metrics (e.g., water quality) may require advanced remote sensing methods or additional datasets (e.g., hyperspectral imagery).

Geographically Weighted Regression (GWR)

$$Y_i = \beta_0 + \sum_{j=1}^{n} \beta_j X_{ij} + \epsilon_i$$

where:

- Y_i = Dependent variable [Remote sensing data, national surveys]
- X_i = Independent variables [Remote sensing data, GIS databases]
- ϵ_i = Error term

Economic Impact Quantification

Objectives of CGE Modeling

- Quantify the economic impacts of ETM extraction and production on sectors such as agriculture, industry, and services.
- Assess intersectoral linkages and shifts in resource allocation driven by mining activities.
- Identify regional disparities and localized economic effects of ETM activities using spatially resolved data.
- Integrate satellite-based metrics with traditional economic data to enhance model accuracy and scalability.

Cross-Country CGE Modeling in Southern Africa

For cross-country CGE modeling in Southern Africa, the model leverages national-level input-output (I/O) tables, trade data, and satellite-derived proxies to refine parameters.

Data Availability for Cross-Country Analysis

1. **National I/O Tables:**
   Sources include national statistics agencies of South Africa, Namibia, Botswana, Zimbabwe, and other Southern African nations. Global sources include the World Input-Output Database (WIOD) and UN Comtrade.

2. **Trade Data:**
   - Traditional sources: UN Comtrade for exports and imports. - Satellite-derived proxies:
     - Night-time lights at port facilities (VIIRS/DMSP).
     - Shipping traffic density from SAR data (e.g., Sentinel-1).

3. **Satellite-Derived Inputs for CGE Models:**
   - **Land use/land cover (LULC) changes:** Proxy for sectoral shifts such as agricultural expansion or mining. Sentinel-2 or MODIS datasets are commonly used for classification.
   - **Economic activity clusters:** Night-time light intensity (VIIRS) to identify high-output regions.
   - **Transportation and infrastructure data:** Road network expansion or density using Sentinel-2 or Planet Labs data.

Limitations Across Countries

- National statistics may not align in methodology or resolution.
- Satellite proxies can help standardize metrics across borders but require validation.

Subnational (District-Level) CGE Modeling

For district-level CGE modeling, where traditional data like I/O tables or trade data is often scarce, satellite-derived metrics are employed to approximate factors of production and sectoral outputs.

Steps for District-Level Analysis

1. **Proxying Factors of Production (X_i):**
   - **Agricultural production:** NDVI/EVI from Sentinel-2 or MODIS as proxies for crop health and yields; CHIRPS rainfall data to model water availability.
   - **Industrial activity:** Surface reflectance anomalies to detect mines and factories (Landsat, Sentinel-2); thermal anomalies to identify industrial zones (MODIS LST).
   - **Human capital:** Population density estimates (WorldPop or GHSL); night-time lights as a proxy for urban workforce availability.
2. **Sectoral Outputs (Y_j):**
   - **Agriculture:** NDVI changes over time to track production trends; crop type classification using Sentinel-2.
   - **Mining and industry:** Surface changes in mining areas (Sentinel-1/2); emission data from Sentinel-5P TROPOMI for NO and SO.
   - **Services:** Night-time lights as a general proxy for service sector growth; urban expansion indicators using built-up indices.
3. **Infrastructure and Transportation (A_{ij}):**
   - **Road density and connectivity:** Extract road networks using Sentinel-2 or OpenStreetMap overlays.
   - **Power infrastructure:** Expansion of energy grids detected via night-time lights or solar panel mapping using high-resolution Sentinel-2 imagery.
   - **Irrigation and water use:** Identify new irrigation systems from NDWI or SAR (Sentinel-1).
4. **Spatially Explicit Model Calibration:**
   - Link satellite proxies to district-level economic metrics.
   - Validate NDVI or night-time lights against known agricultural or industrial outputs in well-documented districts.
   - Use regression models to extrapolate missing data for underserved districts.

Limitations at Subnational Levels

- Scarcity of ground truth data for calibration.
- Spatial resolution of satellite data may not capture fine-grained economic activities.

Appendix: Data Sources and References

- Statistics South Africa: `https://www.statssa.gov.za/`
- National I/O Tables Report: `https://www.statssa.gov.za/publications/Report-04-04-02/Report-04-04-022014.pdf`
- UN Comtrade Database: `https://comtrade.un.org/`
- Night-time lights data: VIIRS/DMSP datasets, accessible via `https://eogdata.mines.edu/`
- Sentinel satellite data: Copernicus Open Access Hub `https://scihub.copernicus.eu/`
- Rainfall data: CHIRPS dataset `https://www.chc.ucsb.edu/data/chirps`
- Land use and vegetation data: MODIS (via Google Earth Engine) `https://developers.google.com/earth-engine/datasets/catalog/MODIS`

Structural Transformation Economic Model

This model evaluates the impact of employment levels (E_t) and industrial investment (I_t) on economic growth as measured by GDP (GDP_t) over time.

Model Framework

Core Equation:

$$\frac{dGDP_t}{dt} = f(E_t, I_t)$$

where:

- GDP_t: Gross Domestic Product at time t, measured in national currency or constant USD.
- E_t: Employment levels at time t, derived from labor force surveys or satellite proxies.
- I_t: Industrial investment at time t, based on national investment data or sector-specific reports.

Functional Form

Assuming a linear approximation for $f(E_t, I_t)$, the equation becomes:

$$\frac{dGDP_t}{dt} = \alpha E_t + \beta I_t$$

where:

- α: Marginal effect of employment on GDP growth.
- β: Marginal effect of industrial investment on GDP growth.

Satellite Integration in the Model

Satellite-derived data can improve estimates for E_t and I_t when traditional data is unavailable or unreliable.

Proxying Employment Levels (E_t)

- **Urban Workforce Proxies:**
  - Night-time light intensity (VIIRS/DMSP): Correlates with urban economic activity and employment.
  - Built-up area indices: Derived from Sentinel-2 or MODIS to estimate urbanization and potential workforce concentrations.
- **Rural Workforce Proxies:**
  - NDVI/EVI (Sentinel-2/MODIS): Reflects agricultural activity, a major employer in rural areas.
  - Population density maps (WorldPop, GHSL): Provides spatial estimates of rural labor distribution.

Proxying Industrial Investment (I_t)

- **Infrastructure Development:**
  - Road network expansion (Sentinel-2, OpenStreetMap).
  - Energy infrastructure expansion, detectable via night-time lights (VIIRS).
- **Mining and Industrial Zones:**
  - Surface reflectance anomalies indicating active mining or industrial activity (Landsat/Sentinel-2).
  - Thermal anomalies from MODIS Land Surface Temperature (LST) data for industrial hotspots.

Numerical Example

For illustrative purposes, assume the following parameter values:

$$\begin{aligned} \alpha &= 0.8, & \beta &= 1.2 \\ E_t &= 500\,(\text{in thousands}), & I_t &= 200\,(\text{in billions of USD}) \end{aligned}$$

The GDP growth rate is:

$$\frac{dGDP_t}{dt} = 0.8 \cdot 500 + 1.2 \cdot 200 = 640\,(\text{in billions of USD per year})$$

Data Sources

- Employment data:
  - National labor force surveys (https://www.statssa.gov.za/).
  - Night-time light proxies (VIIRS, https://eogdata.mines.edu/).
- Industrial investment data:
  - Industry reports (https://unctad.org/).
  - Surface reflectance and thermal anomaly data (Sentinel-2, MODIS).
- GDP data:
  - National accounts (World Bank, https://data.worldbank.org/).
  - Night-time light intensity as a proxy.

Satellite Imagery Analysis

- **Random Forest Classification for Land Use Changes**

$$P(X) = \sum_{t=1}^{T} \left(\sum_{i=1}^{M} w_i f(X_i) \right)$$

  where:

  - $P(X) =$ Probability of land use classification [Landsat, MODIS, Google Earth Engine]
  - $w_i =$ Weights of decision trees [Random Forest algorithm]
  - $f(X_i) =$ Decision function based on spectral features [Landsat, Sentinel-2]

- **NDVI Time Series Analysis**

$$NDVI_t = \frac{NIR - RED}{NIR + RED}$$

  where:

- $NDVI_t$ = NDVI value [MODIS, Sentinel-2, Landsat]
- NIR = Near-infrared band [Satellite imagery]
- RED = Red band [Satellite imagery]

- **Landscape Fragmentation Index Calculation**

$$LFI = \frac{P}{\sqrt{A}}$$

where:

- LFI = Landscape fragmentation index [Landsat, MODIS]
- P = Patch perimeter [Remote sensing data]
- A = Area [Remote sensing data]

Objective 2: Illegal Sand Mining Monitoring Framework

Machine Learning Detection Models

- **Convolutional Neural Network (CNN) Architecture**

$$X_i \to \text{CNN layers} \to \text{Output layer}$$

where:

- X_i = Satellite image [Sentinel-1, Sentinel-2, WorldView]
- y_i = Label for mining activity [Mining datasets, satellite data]

- **U-Net Semantic Segmentation Model**

$$X_i \to \text{U-Net layers} \to Y_i$$

where:

- X_i = Satellite image [WorldView, GeoIQ]
- Y_i = Segmentation map [Mining labeled datasets]

- **Object Detection using YOLO Algorithm**

$$X_i \to \text{YOLO layers} \to \text{Bounding box for detected mining areas}$$

where:

- X_i = Satellite image [High-resolution imagery]
- y_i = Bounding box for detected mining areas [Satellite datasets]

Spectral and Temporal Analysis

- **Multi-temporal Change Detection Algorithm**

$$\Delta X_i = X_{i,t2} - X_{i,t1}$$

where:

  - ΔX_i = Change detection [MODIS, Landsat, Sentinel-2]
  - $X_{i,t2}$ = Image at time t_2
  - $X_{i,t1}$ = Image at time t_1

- **Spectral Mixture Analysis (SMA)**

$$L_i = \sum_{j=1}^{n} a_{ij} f_j$$

where:

  - L_i = Land cover types [Sentinel-2, MODIS, Landsat]
  - a_{ij} = Spectral mixing coefficients [Field data, satellite data]
  - f_j = Spectral endmembers [Satellite data]

- **Normalized Difference Water Index (NDWI) Transformation**

$$NDWI_t = \frac{Green - NIR}{Green + NIR}$$

where:

  - $NDWI_t$ = Water presence index [Sentinel-1, Sentinel-2, MODIS]
  - $Green$ = Green band [Satellite imagery]
  - NIR = Near-infrared band [Satellite imagery]

Objective 3: Socio-Economic Impacts of Sand Extraction

Econometric Modeling

5 Key Questions

Identify the key socio-economic outcomes and issues for analysis. For sand mining, possible research questions include:

- How does sand mining affect GDP, employment, and trade in sand-dependent regions?

- What are the environmental spillovers (e.g., changes in vegetation, water quality) and their subsequent impact on local economies?
- What is the lagged effect of sand mining on infrastructure development and population migration?

6 Formulating Hypotheses

Develop hypotheses about causal relationships between sand mining and socio-economic outcomes:

- **H1:** Increased sand extraction leads to short-term GDP growth in affected regions.
- **H2:** Sand mining causes environmental degradation, negatively impacting agriculture and employment over time.
- **H3:** Infrastructure expansion driven by sand mining activities spurs urbanization.

7 Panel VAR Model Specification

The generic Panel VAR model is specified as:

$$Y_{jt} = \alpha_j + \sum_{i=1}^{n} \beta_i X_{jit} + \sum_{k=1}^{p} \gamma_k Y_{j(t-k)} + \epsilon_{jt} \tag{1}$$

Where:

- Y_{jt}: Socio-economic outcomes (e.g., GDP, employment, migration patterns) for region j at time t.
- X_{jit}: Sand mining and environmental indicators (e.g., sand extraction rates, NDVI, water turbidity) for region j at time t.
- $Y_{j(t-k)}$: Lagged values of socio-economic outcomes (e.g., previous GDP or employment levels).
- α_j: Region-specific fixed effects.
- β_i, γ_k: Coefficients to estimate relationships.
- ϵ_{jt}: Error term.

8 Data and Variables

8.1 Dependent Variables (Y_t)

Economic outcomes:

- GDP growth (Y_{GDP}): Data from World Bank, IMF, or national statistics.
- Employment in affected sectors (Y_{EMP}): Data from ILO or labor surveys.
- Trade metrics (Y_{TRADE}): Export and import values related to construction materials from UN Comtrade.

Social outcomes:

- Migration patterns (Y_{MIG}): Data from WorldPop or census data.
- Urbanization rates (Y_{URB}): Satellite-derived urbanization data (e.g., Sentinel-2, Landsat).

8.2 Independent Variables (X_t)

Sand mining activity:

- Extraction volume: Local permits, industry reports.
- Satellite proxies: Surface reflectance anomalies, deforestation near mining sites.

Environmental indicators:

- Vegetation health (NDVI): Data from Sentinel-2 or MODIS.
- Water turbidity: Data from Sentinel-2 MSI or MODIS Aqua.

8.3 Control Variables (C_t)

- Rainfall: Data from CHIRPS or IMERG.
- Infrastructure development: Data from OpenStreetMap or government reports.
- Economic policies: National reports on mining regulations.

Spatial Durbin Model for Coastal Community Effects

The Spatial Durbin Model (SDM) captures spatial dependencies in both the dependent and independent variables, making it suitable for studying spillover effects in coastal communities.

$$Y_i = \rho W_i Y + \beta X_i + \epsilon_i$$

Where:

[noitemsep]Y_i: Community-level outcomes (e.g., socio-economic and environmental effects). W_i: Spatial weight matrix (e.g., adjacency or distance-based). X_i: Independent variables (e.g., socio-economic and environmental indicators). ρ: Coefficient for the spatial lag of the dependent variable. β: Coefficients for the independent variables. ϵ_i: Error term.

Variables and Data Sources

Dependent Variable (Y_i): Community Effects

[noitemsep]**Remote Sensing Data:**[noitemsep]

- - Vegetation health (NDVI): MODIS, Sentinel-2.
  - Water quality (turbidity, chlorophyll): Sentinel-2 MSI, Landsat.
- **Socio-Economic Data:**
  [noitemsep]Employment rates, migration: Census data, WorldPop. GDP growth: World Bank, IMF.

Spatial Weight Matrix (W_i)

[noitemsep]**Community Data:**[noitemsep]

- - Coastal community locations: OpenStreetMap, national censuses.
  - Population density: WorldPop.
- **Industrial Data:**
  [noitemsep]Industry locations: UNIDO, national industrial registries. Infrastructure networks: OpenStreetMap.

Independent Variables (X_i)

[noitemsep]Socio-economic indicators: Census data, World Bank. Environmental indicators (NDVI, water quality): Sentinel-2, MODIS.

Propensity Score Matching for Causal Inference

Propensity Score Matching (PSM) estimates causal effects by comparing treated and control units with similar covariates.

$$Y_i = \beta_0 + \beta_1 T_i + \epsilon_i$$

Where:

[noitemsep]Y_i: Outcome measure (e.g., socio-economic or environmental effects). T_i: Treatment indicator (e.g., industrial exposure). β_0, β_1: Coefficients to be estimated. ϵ_i: Error term.

Variables and Data Sources

Outcome Measure (Y_i)

[noitemsep]**Socio-Economic Data:**[noitemsep]

- – Income levels, employment rates: Census data, labor force surveys.
  – Migration rates: WorldPop, UN Population Division.
  – Urbanization: Landsat, Sentinel-2.
- **Environmental Data:**
  [noitemsep]Vegetation health (NDVI): MODIS, Sentinel-2. Water quality (turbidity): Sentinel-2 MSI, Landsat.

Treatment Indicator (T_i)

[noitemsep]Industrial exposure:[noitemsep]

- – Proximity to sand mining or industrial sites: UNIDO, OpenStreetMap.
  – Environmental degradation: MODIS, Sentinel-2.
- Policy exposure: National mining regulations, industry-specific policies.

Matching Variables (Covariates)

[noitemsep]Baseline socio-economic data: Household income, education levels (World Bank, national surveys). Environmental conditions: NDVI, water quality metrics before industrial activity (MODIS, Sentinel-2).

Tools for Both Models

[noitemsep]**Software:**[noitemsep]

- - Python: `PySAL` (spatial models), `statsmodels` (PSM).
  - R: `spdep` (spatial analysis), `MatchIt` (PSM).
  - GIS Platforms: ArcGIS, QGIS.
- **Data Sources:**
  [noitemsep]Socio-economic data: World Bank, IMF, UN Comtrade. Environmental data: Sentinel-2, MODIS, Landsat. Spatial data: OpenStreetMap, national registries.

Objective 4: Quantifying Industrial Environmental Costs

Environmental Cost Estimation Models

- **Integrated Environmental-Economic Accounting Framework**

$$C_e = \sum_{i=1}^{n} E_i \quad \text{where} \quad C_e = \text{Environmental cost}, E_i = \text{Environmental impact}$$

  where:

  - C_e: Environmental cost [Environmental-economic data, industry reports]
  - E_i: Environmental impact [Industry reports, satellite data]
- **Spatial Hedonic Pricing Model**

$$H_i = \alpha + \beta_1 E_i + \beta_2 D_i + \epsilon_i$$

  where:

  - H_i: Housing price [Real estate data, proximity to industries]
  - E_i: Environmental quality [Air quality, water quality]
  - D_i: Distance to pollution source [Proximity data]
- **Generalized Method of Moments (GMM) Estimation for Industrial Externalities**

$$E_i = \alpha + \beta_1 X_i + \epsilon_i$$

  where:

  - E_i: Environmental externalities [Satellite data, economic impact data]
  - X_i: Industrial activity [Industry data]

- **Socioeconomic Impact Model for Industrial Environmental Costs**

  $$Y_i = \gamma_0 + \gamma_1 C_e + \gamma_2 S_i + \epsilon_i$$

  where:

  - Y_i: Socioeconomic outcomes (e.g., income, employment, health indicators) [Census data, house
  - C_e: Environmental cost [Environmental-economic data]
  - S_i: Socio-demographic factors (e.g., education, age, occupation) [Census data, labor force surv
  - $\gamma_0, \gamma_1, \gamma_2$: Coefficients to estimate.
  - ϵ_i: Error term.

Data Sources for Models

- **Environmental Data:**
  - Satellite data (NDVI, air quality, water turbidity): Sentinel-2, MODIS, Landsat.
  - Industry reports and permits: Local and national environmental agencies.
- **Economic Data:**
  - GDP, employment rates, household income: World Bank, IMF, national statistics.
  - Real estate data: National real estate registries, market surveys.
- **Socioeconomic Data:**
  - Census data, labor force surveys, migration: WorldPop, ILO, national statistics.