Sand Mining Watch: Leveraging Earth Observation Foundation Models to Inform Sustainable Development

Ando Shah
School of Information
University of California, Berkeley
ando@berkeley.edu

Suraj R. Nair
School of Information
University of California, Berkeley
suraj.nair@berkeley.edu

Tom Böhnel
Environmental Sensing and Modeling
Technical University of Munich
tom.boehnel@tum.de

Joshua E. Blumenstock
School of Information
University of California, Berkeley
jblumenstock@berkeley.edu

Abstract

As the major ingredient of concrete and asphalt, sand is vital to economic growth, and will play a key role in aiding the transition to a low carbon society. However, excessive and unregulated sand mining in the Global South has high socio-economic and environmental costs, and amplifies the effects of climate change. Sand mines are characterized by informality and high temporal variability, and data on the location and extent of these mines tends to be sparse. We propose to build custom sand-mine detection tools by fine-tuning foundation models for earth observation, which leverage self supervised learning - a cost-effective and powerful approach in sparse data regimes. Our preliminary results show that these methods outperform fully supervised approaches, with the best performing model achieving an average precision score of 0.57 for this challenging task. These tools allow for real-time monitoring of sand mining activity and can enable more effective policy and regulation, to inform sustainable development.

1 Introduction

Sustainable sand mining is one of the main ecological challenges of the 21st century[1–3]. Driven by rapid urbanization and economic growth[1], sand mining activity has tripled over the last 20 years, and accounts for 85% of all mineral extraction[4]. The demand for sand is expected to increase many-fold, given the key role of sand-intensive infrastructure in aiding the transition to a low-carbon society[2] [5] and development. Across the Global South, much of this sand is mined from riverbeds and riverbanks, and has acute socio-economic and environmental costs. In particular, it threatens to amplify the impacts of climate change by accelerating flooding, erosion and biodiversity loss[1, 6]. Despite these adverse consequences, sand mining remains nearly unregulated in most parts of the world, especially in India where the rate of extraction already far exceeds sustainable levels [7–9]. Thus, urgent action is vital to prevent a global sand crisis [10].

The highly pervasive nature of sand mining makes physical monitoring near impossible, given the scale of the task and the dangers involved [2, 11]. While the use of satellite imagery to detect

[1]Sand constitutes 70-85% of concrete and asphalt, critical for construction, roads, and other infrastructure
[2]Demand is expected to increase from municipal infrastructure, utilities, social infrastructure and buildings.

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2023.

mining activity is common [12–15], sand mining is characterized by high amounts of informality and temporal variability, and is often hard to distinguish from natural fluvial processes. As a result, sand mines are commonly unrepresented in general purpose mine detection tools, most of which focus on metal extraction.

We propose to address these challenges by using self-supervised learning (SSL), in combination with open-access medium-resolution satellite imagery, and a new dataset of hand-crafted labels to build a custom sand mine detection tool. The overarching goal of our project is to produce, and enable the continued production of high-resolution, real-time maps of sand mining activity. Thus far, a small body of literature has examined the use of satellite imagery and machine learning to map global distributions of sand deposits [16–19]. This project expands on these ideas, and contributes to the research on the use of satellite imagery and machine learning to monitor mining activity [12, 14, 20]. In particular, we leverage advances in SSL methods that have recently seen success in the context of earth observation (EO) [21–26]. Here, a global corpus of unlabeled multispectral remote sensing datasets are used to learn semantic representations of the planet's surface from freely available satellite imagery. These pre-trained models, often referred to as *foundation models for earth observation* (FMEO), have shown impressive performance on downstream tasks such as object detection and image segmentation [23]. Most importantly, FMEOs typically exploit all available multi-spectral bands of optical imagery, and some additionally use synthetic aperture radar (SAR) data. From existing work on mapping sand and gravel using remote sensing [16, 18, 27–29], we know that RGB, shortwave infrared, thermal bands, and SAR comprise the bands of importance, underscoring the importance of FMEOs in this transfer learning paradigm. Our approach will test the ability of these models to generalize to sand mine segmentation, a novel and perhaps more challenging task than these models have been tested on previously. In the process we build an understanding of the best practices required to fine tune these models, especially under distribution shifts. In our analysis, we compare and contrast the performance of FMEOs with that of fully-supervised approaches.

2 Data and Methods

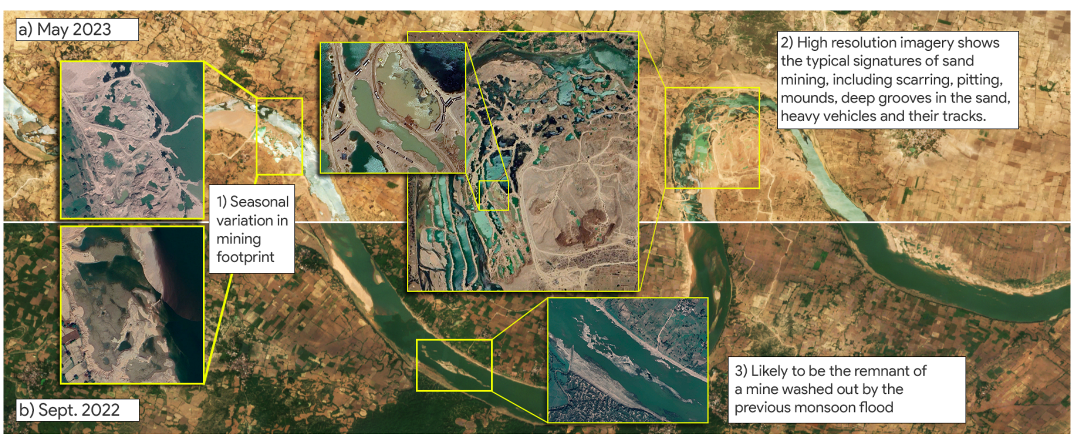

Figure 1: Mining signatures and seasonal variability: Both panels show a section of the river Betwa, in the state of Uttar Pradesh, India. Panel a) and b) are monthly median composites of 10m Sentinel-2 L2A RGB images taken in May 2023 and September 2022, respectively. Insets 1 - 3 contain sub-meter resolution imagery, and describe some of the nuances and peculiarities of this detection task.

Datasets: We have acquired data (latitude, longitude, timestamp) on sand mining activities across 21 different river basins across India, through a partnership with Veditum India Foundation[3]. Currently, these data cover 39 distinct mining sites; we expect to expand this to over 100 sites over the course of our study. We extract arbitrary-sized observations (ranging in size from 2.5 sq.km to 582 sq.km) from freely available Sentinel-2 multi-spectral and Sentinel-1 synthetic aperture radar imagery

[3] India Sand Watch - https://sandmining.envmonitoring.in

Model	Average Precision [%]	F1 Score [%]	Input Sentinel bands	Strategy	Trainable / Total parameters
U-Net	50.75	48.84	6 L2A bands	Fully supervised	31M / 31M (100%)
SegFormer-B0	31.60	41.80	6 L2A bands	Fully supervised	4M / 4M (100%)
SSL4EO-Resnet18*	47.93	47.86	13 L1C bands	Linear probe	3M / 14M (22%)
SSL4EO-Resnet50*	**57.11**	49.26	13 L1C bands	Linear probe	48M / 72M (67%)
SatMAE-base*	49.36	48.41	10 L2A bands	Linear probe	0.8M / 86M (0.9%)
SatMAE-large*	55.82	**53.27**	10 L2A bands	Linear probe	1M / 304M (0.3%)

Table 1: Results on the spatially held-out test set. Performance metrics are denoted in percentage points. * indicates pretrained models. Entries in bold are the best performing models for AP and F1 scores.

around visually recognizable sand mining footprints at each site[4]. A majority of Indian rivers are characterized by high average flood discharges and large temporal variability [30], leading to huge intra-annual variation in sand deposition rates and mining footprints. We consider these changes to be strong natural label augmentations (Fig. 1, inset 1). This allows us to obtain multiple labels (of arbitrary size) for each location that represents the seasonal lifecycle of sand mines. While sub-meter resolution imagery (Fig. 1, inset 2) captures more precise information on mining activity such as direct evidence of heavy machinery, we believe that the Sentinel 10m imagery will prove to be an effective feature set since it captures broad patterns of importance (i.e. scarring, pitting and flooding) at high temporal & spectral resolution, and is freely available.

Methods: We use river extent polygons downloaded from Open Street Maps to create regions of interest extending up to 1km on either side of given rivers. We randomly samples image patches of size 160x160 pixels (1.6km x 1.6km) during training, and use a sliding window to sample the entire region of interest during validation and inference. Using this method, we generate ~ 3100 images for training per epoch and 380 images for validation. Given the low-data regime this problem inhabits, we propose to leverage SSL methods, specifically the use of pre-trained FMEOs and fine-tune them to solve the task of sand mine segmentation. We hypothesize that our approach will lead to higher performance compared to fully-supervised methods, due to limited learning that is possible with a small set of labels in a fully-supervised setting. Specifically, we will use SatMAE[21], SSL4EO-S12[22], Satlas[25] and Scale-MAE [26] models [5] We compare the performance of these models to fully supervised CNN-based models including U-Nets [33] and SegFormer [34].

We frame this problem as a binary semantic segmentation task with a focus on characterizing the performance of the detection of the sand mine class. Our labels are highly imbalanced with the "sandmine" class comprising only 6.07% of the total number of pixels (the remaining being "non-sandmine"). We apply an inverse-weighting scheme to the binary cross entropy loss during training in order to account for this imbalance. Both our labels and features are spatially autocorrelated; we thus cluster observations spatially, and train and validate separately on geographically distinct regions across India. This helps us to avoid known challenges with standard evaluation approaches, which tend to overestimate performance in settings with spatial autocorrelation [35, 36].

In addition to conventional fine-tuning of the FMEOs, we will also explore methods to ameliorate the possible distribution shift issues incurred by transfer learning. This is especially of concern with the pre-training datasets used by SatMAE (fMoW-Sentinel) and SSL4EO (SSL4EO-S12); the former exhibits a global north bias and the latter, a strong urban bias - by virtue of their sampling methodologies. Moreover, the geographical areas of interest for this problem span a very specific domain - river basins of South Asia, which is likely to be highly under-represented in the pre-training datasets. We propose to use methods like test-time training[37], which is a form of transductive learning[38] whereby, for each test image, a local model is trained by self-supervision. An extreme form of this would be to learn unsupervised representations of the entire test/validation dataset, which may also produce desirable results.

[4]Labelers examine sites using medium-resolution Sentinel-2 RGB imagery, high resolution imagery available from Google Earth Pro (Maxar, 0.3m) and from Norway's International Climate and Forests Initiative (NICFI) program (Planet, 4.5m), when available.

[5]These rely on self-supervised approaches that include masked autoencoders (MAE)[31], DINO [32] and MoCo[32] for both CNN and ViT-based architectures.

3 Preliminary Results

In policy settings, it is often desirable to balance false positives and false negatives, and therefore F1 scores are typically employed as the primary metric. However, instead of manually choosing a trade-off between precision and recall, we use the average precision (AP) metric to summarize the area under the entire precision-recall curve, which is also less sensitive to class imbalance exhibited by our dataset. Early experiments with short training runs (20-40 epochs, see Table 1) show that foundation models outperform fully-supervised methods. Our best performing model (SSL4EO-ResNet50) exhibits an AP of 0.57 for the sand mine class. However, the F1 score for this model is 0.49 which comprises a relatively high recall (0.69) and low precision (0.39). To characterize the performance further, we conducted qualitative analyses on the test set and Fig. 2 shows an instance of the prediction. Fig. 2 panel a) shows ground-truth labels (blue) and predictions (red) by the best performing model; areas in purple show overlap. The yellow ellipses highlight labels with medium-to-high uncertainty; indeed, we observe poor model performance in these regions. Currently, even though human annotators experienced uncertainty about mine extents, we pursued a binary labeling strategy (mine/no-mine). To ameliorate this issue moving forward, we are examining the use of multi-class labels which capture different levels of uncertainty during the labeling process. With this new labeling strategy, along with expanding the training dataset and distribution shift amelioration methods during fine-tuning, we expect substantial performance improvements in the near future.

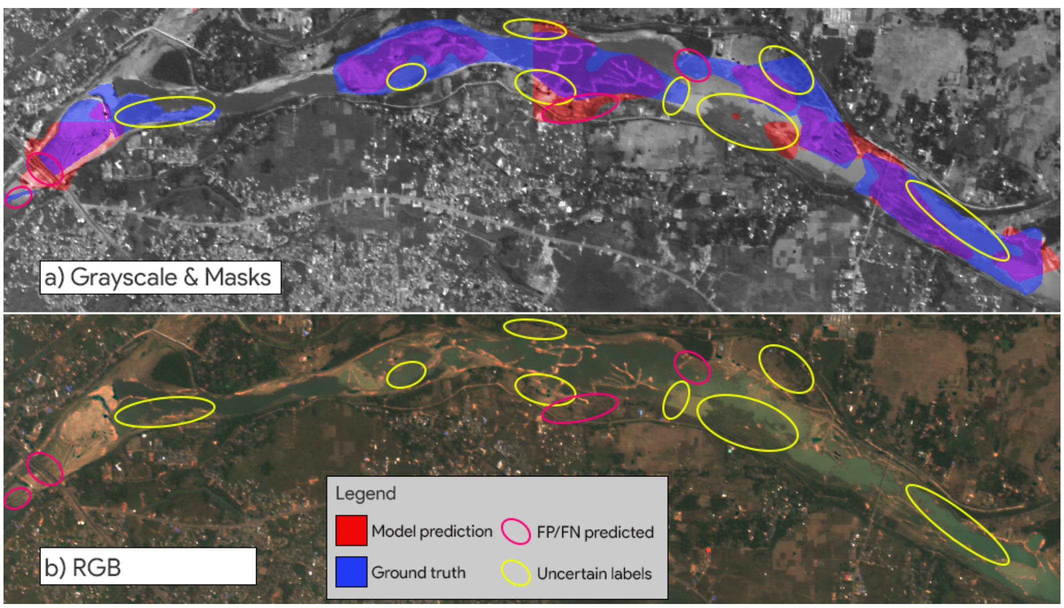

Figure 2: Qualitative assessment of predictions compared to labels. Panels a) & b) show the same section of the river Bhargavi in the state of Odisha, India using Sentinel-2 imagery, grayscale and RGB, respectively.

4 Pathways to Climate Impact

This work aims to provide policymakers and civil society with the tools to monitor and regulate sand mining in an effective and sustainable way, moving closer to managing and mitigating the harmful effects of climate change on river ecosystems. Our work also furthers progress towards Sustainable Development Goals (SDG) related to environmental protection on land (SDG 15), responsible consumption and production (SDG 12), clean water and sanitation (SDG 6), and sustainable cities and communities (SDG 11). Finally, the use of FMEO allows researchers to produce better performing models at reduced computational cost and carbon footprint by leveraging pre-trained models. By conducting these downstream experiments, we contribute to the understanding of the nuances of operationalizing such models, and incentivize model reuse in earth observation problems.

References

[1] Aurora Torres, Mark U. Simoni, Jakob K. Keiding, Daniel B. Müller, Sophus O. S. E. zu Ermgassen, Jianguo Liu, Jochen A. G. Jaeger, Marten Winter, and Eric F. Lambin. Sustainability of the global sand system in the Anthropocene. *One Earth*, 4(5):639–650, May 2021. ISSN 2590-3322. doi: 10.1016/j.oneear.2021.04.011.

[2] Paul Salopek. Inside the deadly world of India's sand mining mafia. *National Geographic*, June 2019.

[3] Pascal Peduzzi, Zhanyl Moldalieva, Louise Gallagher, Charlotte Poussin, and Davide Fornacca. Sand and sustainability UNEP 2019, 2019.

[4] Fred Pearce. The Hidden Environmental Toll of Mining the World's Sand. *Yale Environment 360*.

[5] Dimitra Ioannidou, Guido Sonnemann, and Sangwon Suh. Do we have enough natural sand for low-carbon infrastructure? *Journal of Industrial Ecology*, 24(5):1004–1015, 2020. ISSN 1530-9290. doi: 10.1111/jiec.13004.

[6] Lois Koehnken and Max Rintoul. Uncovering sand mining's impacts on the world's rivers. *WWF*, 2018.

[7] Arnaud Vander Velpen, Aurora Torres, and Pascal Peduzzi. Sand and Sustainability Report. Technical report, UNEP, 2022.

[8] Monika Mondal. The Indian river being hollowed by mining for concrete. *The Third Pole*, February 2022.

[9] Marco Hernandez, Katy Daigle, and Simon Scarr. The messy business of sand mining explained. *Reuters*, 2021.

[10] UNEP/GRID-Geneva. What is sand? Result from a UNEP/GRID-Geneva expert discussion. 2022. doi: 10.13097/ARCHIVE-OUVERTE/UNIGE:160291.

[11] Rollo Romig. How to Steal a River. *The New York Times*, March 2017. ISSN 0362-4331.

[12] Jane Gallwey, Carlo Robiati, John Coggan, Declan Vogt, and Matthew Eyre. 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, October 2020. ISSN 0034-4257. doi: 10.1016/j.rse.2020.111970.

[13] Detecting illegal mines from space. https://dida.do/blog/detecting-illegal-mines-from-space.

[14] Mathieu Couttenier, Sebastien Di Rollo, Louise Inguere, Mathis Mohand, and Lukas Schmidt. Mapping artisanal and small-scale mines at large scale from space with deep learning. *PLoS ONE*, 17(9):e0267963, September 2022. ISSN 1932-6203. doi: 10.1371/journal.pone.0267963.

[15] Liang Tang and Tim T. Werner. Global mining footprint mapped from high-resolution satellite imagery. *Communications Earth & Environment*, 4(1):1–12, April 2023. ISSN 2662-4435. doi: 10.1038/s43247-023-00805-6.

[16] Ando Shah and Suraj R. Nair. Mapping Construction Grade Sand: Stepping Stones Towards Sustainable Development. In *2023 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Workshop: Fragile Earth: AI for Climate Sustainability - from Wildfire Disaster Management to Public Health and Beyond*, July 2023.

[17] Arjen Luijendijk and Sierd de Vries. 26 - Global beach database. In Derek W. T. Jackson and Andrew D. Short, editors, *Sandy Beach Morphodynamics*, pages 641–658. Elsevier, January 2020. ISBN 978-0-08-102927-5. doi: 10.1016/B978-0-08-102927-5.00026-6.

[18] Giulia Marchetti, Simone Bizzi, Barbara Belletti, Barbara Lastoria, Francesco Comiti, and Patrice Enrique Carbonneau. Mapping riverbed sediment size from Sentinel-2 satellite data. *Earth Surface Processes and Landforms*, 47(10):2544–2559, 2022. ISSN 1096-9837. doi: 10.1002/esp.5394.

[19] Kevin K Williams and Ronald Greeley. Laboratory and field measurements of the modification of radar backscatter by sand. *Remote Sensing of Environment*, 89(1):29–40, January 2004. ISSN 00344257. doi: 10.1016/j.rse.2003.09.006.

[20] Santiago Saavedra. Technology and State Capacity: Experimental Evidence from Illegal Mining in Colombia. *SSRN*, April 2023. doi: 10.2139/ssrn.3933128.

[21] Yezhen Cong, Samar Khanna, Chenlin Meng, Patrick Liu, Erik Rozi, Yutong He, Marshall Burke, David B. Lobell, and Stefano Ermon. SatMAE: Pre-training Transformers for Temporal and Multi-Spectral Satellite Imagery, January 2023.

[22] Yi Wang, Nassim Ait Ali Braham, Zhitong Xiong, Chenying Liu, Conrad M. Albrecht, and Xiao Xiang Zhu. SSL4EO-S12: A Large-Scale Multi-Modal, Multi-Temporal Dataset for Self-Supervised Learning in Earth Observation, May 2023.

[23] Alexandre Lacoste, Nils Lehmann, Pau Rodriguez, Evan David Sherwin, Hannah Kerner, Björn Lütjens, Jeremy Andrew Irvin, David Dao, Hamed Alemohammad, Alexandre Drouin, Mehmet Gunturkun, Gabriel Huang, David Vazquez, Dava Newman, Yoshua Bengio, Stefano Ermon, and Xiao Xiang Zhu. GEO-Bench: Toward Foundation Models for Earth Monitoring, June 2023.

[24] Keumgang Cha, Junghoon Seo, and Taekyung Lee. A Billion-scale Foundation Model for Remote Sensing Images, April 2023.

[25] Favyen Bastani, Piper Wolters, Ritwik Gupta, Joe Ferdinando, and Aniruddha Kembhavi. SatlasPretrain: A Large-Scale Dataset for Remote Sensing Image Understanding, August 2023.

[26] Colorado J. Reed, Ritwik Gupta, Shufan Li, Sarah Brockman, Christopher Funk, Brian Clipp, Kurt Keutzer, Salvatore Candido, Matt Uyttendaele, and Trevor Darrell. Scale-MAE: A Scale-Aware Masked Autoencoder for Multiscale Geospatial Representation Learning, April 2023.

[27] Jasper Knight and Mohamed A. M. Abd Elbasit. Characterisation of Coastal Sediment Properties from Spectral Reflectance Data. *Applied Sciences*, 12(13):6826, July 2022. ISSN 2076-3417. doi: 10.3390/app12136826.

[28] D van der Wal and P.M.J. Herman. Quantifying the particle size of intertidal sediments with satellite remote sensing in the visible light, thermal infrared and and microwave spectral domain, 2006.

[29] Benjamin Purinton and Bodo Bookhagen. Multiband (X, C, L) radar amplitude analysis for a mixed sand- and gravel-bed river in the eastern Central Andes. *Remote Sensing of Environment*, 246:111799, September 2020. ISSN 0034-4257. doi: 10.1016/j.rse.2020.111799.

[30] Vishwas S. Kale. Geomorphic Effects of Monsoon Floods on Indian Rivers. In M. Monirul Qader Mirza, Ajaya Dixit, and Ainun Nishat, editors, *Flood Problem and Management in South Asia*, pages 65–84. Springer Netherlands, Dordrecht, 2003. ISBN 978-94-017-0137-2. doi: 10.1007/978-94-017-0137-2_3.

[31] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. Masked Autoencoders Are Scalable Vision Learners, December 2021.

[32] Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging Properties in Self-Supervised Vision Transformers, May 2021.

[33] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation, May 2015.

[34] Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, and Ping Luo. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers, October 2021.

[35] Esther Rolf. Evaluation Challenges for Geospatial ML, March 2023.

[36] Guanghua Chi, Han Fang, Sourav Chatterjee, and Joshua E Blumenstock. Microestimates of wealth for all low-and middle-income countries. *Proceedings of the National Academy of Sciences*, 119(3):e2113658119, 2022.

[37] Yossi Gandelsman, Yu Sun, Xinlei Chen, and Alexei A Efros. Test-Time Training with Masked Autoencoders.

[38] Vladimir Vapnik. *Estimation of Dependences Based on Empirical Data*. Springer Science & Business Media, 2006.